

STIC-Biotech/ChemLib

65994

From: Saoud, Christine
Sent: Monday, May 06, 2002 10:30 AM
To: STIC-Biotech/ChemLib
Subject: sequence search

10/082902

Please search SEQ ID NO:1 and 2 in the patent and commercial databases. Please include interference search (pending).

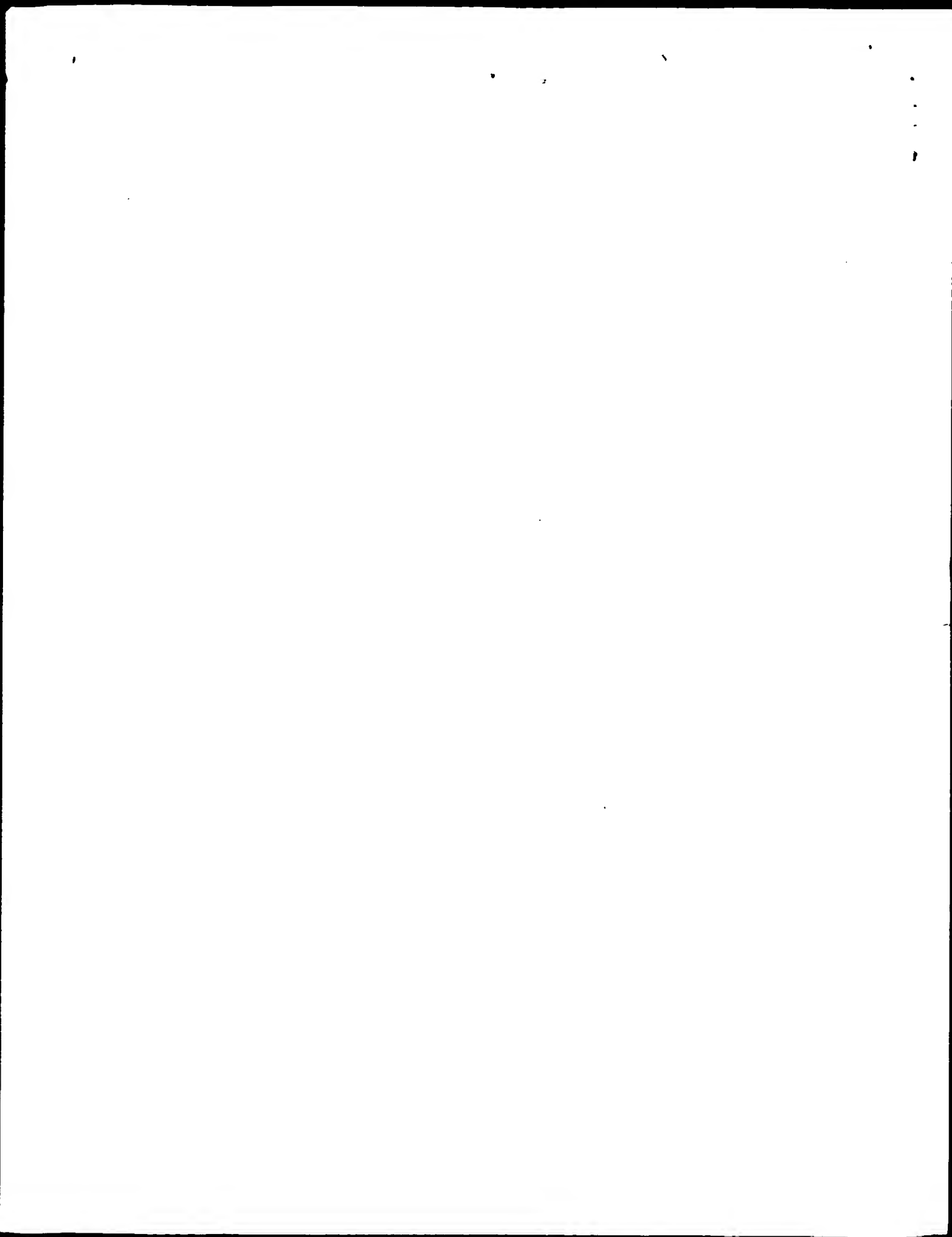
Thanks,
Christine Saoud
A.U. 1647
CM1 - 10E03
305-7519
mailbox 10C01

RECEIVED
MAY - 6 2002
(STIC)

of Contact:
Searcher: Sheppard
Phone: _____
Location: tel: 308-4499
Date Picked Up: _____
Date Completed: 5/11/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2002, 01:06:40 ; Search time 52.66 Seconds
(without alignments)
7752.428 Million cell updates/sec

Title: US-10-082-902-1

Perfect score: 1662
Sequence: 1 agtagcgagctgagctgac.....tgctctggggaagtgtaga 1662

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA: *
1: /cgn2-6/prodata/2/1na/5A_COMB.seq: *
2: /cgn2-6/prodata/2/1na/5B_COMB.seq: *
3: /cgn2-6/prodata/2/1na/6A_COMB.seq: *
4: /cgn2-6/prodata/2/1na/6B_COMB.seq: *
5: /cgn2-6/prodata/2/1na/PCITUS_COMB.seq: *
6: /cgn2-6/prodata/2/1na/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	308.6	18.6	456	US-08-965-903B-15	Sequence 15, App1
2	180.2	10.8	2135	US-08-965-903B-7	Sequence 7, App1
3	173	10.4	1604	US-08-965-903B-5	Sequence 5, App1
4	112.8	6.8	252	US-08-965-903B-12	Sequence 12, App1
5	98.8	5.9	3254	US-08-965-903B-1	Sequence 1, App1
6	97.2	5.8	372	US-08-965-903B-3	Sequence 3, App1
7	64.8	3.9	300	US-08-965-903B-10	Sequence 10, App1
8	52	3.1	7218	US-08-232-463-14	Sequence 14, App1
9	51.4	3.1	3468	US-07-951-715A-2	Sequence 2, App1
10	51.4	3.1	3468	US-08-459-448A-2	Sequence 2, App1
11	51.4	3.1	3468	US-08-459-595A-2	Sequence 2, App1
12	51.4	3.1	3468	US-08-459-504B-2	Sequence 2, App1
13	51.4	3.1	3468	US-08-459-444-2	Sequence 2, App1
14	51.4	3.1	3468	US-09-053-549-3	Sequence 3, App1
15	51.4	3.1	3468	US-09-547-422-2	Sequence 2, App1
16	44.4	2.7	1231	US-09-547-422-2	Sequence 2, App1
17	44.4	2.7	1879	US-09-547-422-2	Sequence 2, App1
18	41	2.5	846	US-07-901-707-12	Sequence 12, App1
19	41	2.5	846	US-07-988-430-12	Sequence 12, App1
20	41	2.5	846	US-08-425-336-12	Sequence 12, App1
21	41	2.5	846	US-08-488-113B-12	Sequence 12, App1
22	41	2.5	846	US-08-477-484B-12	Sequence 12, App1
23	41	2.5	846	US-08-466-360-12	Sequence 12, App1
24	41	2.5	846	US-08-839-765-12	Sequence 12, App1
25	41	2.5	846	US-09-136-388-12	Sequence 12, App1
26	41	2.5	846	PCT-US92-09487-12	Sequence 12, App1
27	40	2.4	800	PCT-US95-04801-4	Sequence 4, App1

c 28	39.8	2.4	2580	3	US-09-050-863-2	Sequence 2, App1
c 29	39.8	2.4	2580	4	US-09-359-081-2	Sequence 2, App1
c 30	39.8	2.4	5452	2	US-09-130-114-1	Sequence 1, App1
c 31	39.8	2.4	9600	4	US-08-910-647-1	Sequence 1, App1
c 32	39.8	2.4	10596	1	US-07-884-811-15	Sequence 15, App1
c 33	39.8	2.4	10596	1	US-07-885-971-15	Sequence 15, App1
c 34	39.8	2.4	10596	1	US-08-087-783A-15	Sequence 15, App1
c 35	39.8	2.4	10596	1	US-08-194-088B-15	Sequence 15, App1
c 36	39.8	2.4	10596	2	US-08-194-087-15	Sequence 15, App1
c 37	39.8	2.4	10596	5	PCT-US93-04648-15	Sequence 15, App1
c 38	39.4	2.4	1032	1	US-08-375-186-1	Sequence 1, App1
c 39	39.4	2.4	1032	1	US-08-457-787A-4	Sequence 4, App1
c 40	39.4	2.4	1032	1	US-08-812-025-4	Sequence 4, App1
c 41	39.4	2.4	1032	4	US-09-138-873A-4	Sequence 4, App1
c 42	39.4	2.4	1032	4	US-08-775-362-1	Sequence 1, App1
c 43	39.2	2.4	4154	1	US-08-131-365B-37	Sequence 37, App1
c 44	39.2	2.4	4154	2	US-08-668-123-37	Sequence 37, App1
c 45	38.4	2.3	9551	1	US-08-056-200-93	Sequence 93, App1

ALIGNMENTS

RESULT 1
US-08-965-903B-15
Sequence 15, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachem, Nir
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SEQUENCE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Petithory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: B600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: m-spry4
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 3...353
OTHER INFORMATION:

US-08-965-903B-15

Query Match	18.6%;	Score 308.6;	DB 3;	Length 456;
Best Local Similarity	86.3%;	Pred. No. 2e-69;		
Matches 341;	Conservative	0;	Mismatches 54;	Indels 0;
			Gaps	0;

QY	755	gcaacacagagatgacctgctgctgctcagccccagacctctgtcctaacctatggacacgtgcatgtttt	814
Db	1	GCACCAGGAGGCGCTGCGCTCGGCTCAGACCGTGGTCACTATGGCATTGCACTGATGTGTC	60
QY	815	tggctgaaggagcattctctaacacactgcagaaatgagagaaatgagggctctctgcgtgcac	874
Db	61	TGGTGAAGAGATCTCTCTCTATCACTGTAATTAATGAGAAATGAGAGGCTCTTGGCCGGAAC	120
QY	875	accctctgctctgctctcccgctccaacactgctgtagccgctctggctcctctatgtagtgcctct	934
Db	121	ACCCGCTGCTCCTGTTCGCGGCTCCAACTGCTGCGCGCGCTGAGCTTCATGAGGGGCCCTCT	180
QY	935	cctgtgtgtctgcctgcctgcctgctcgtcctactcctcctgacacacgctgtgctgaactctggcc	994
Db	181	CTGTGTGCTGTGCGCTGTCTGCTGTGCTGCTACTGTGCGGCTCCACAGGCTGCGTCAAGCTGACCC	240
QY	995	agcgtgtgctacgaacccgtctctgcgcgcgcctgtgtctgcgcctgtaaaagacacagaaagagctca	1054
Db	241	AGCGAGCGCTACACACCGGCTTGAGACGCCGCCGCTTCCGCTGCAACACACGAACAGCGGTA	300
QY	1055	tctgcgaagacagccagcagcgagatgacacagacagacagccagacacacccctttcttgacgt	1114
Db	301	TCTGCAAGCGCAAGCAAGTGGGAGAACCAAGCAAGCAGCGAGCTGTACACACCTTTCTGACACT	360
QY	1115	tgtgtctgcgaagccacagtgctctcgtcctgtgaaacct	1149
Db	361	TGTGATGATAAATAATGATGTGCTCCCTCTGCGCAAGCT	395

RESULT 2
 US-08-965-903B-7
 : Sequence 7, Application US/08965903B
 Patient No. 6060275
 GENERAL INFORMATION:
 APPLICANT: Hachon, Nlr
 APPLICANT: Krasnow, Mark A.
 TITLE OF INVENTION: SPROUTY PROTEIN AND CODING SEQUENCE
 TITLE OF INVENTION: SEQUENCE
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 350 Cambridge Ave., Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/965,903B
 FILING DATE: 07-NOV-1997
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 60/030232
 FILING DATE: 07-NOV-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Petichory, Joanne R
 REGISTRATION NUMBER: 42,995
 REFERENCE/DOCKET NUMBER: 8600-0177, 30
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-0880
 TELEFAX: 650-324-0960
 TELEX:

```

? INFORMATION FOR SEQ. ID NO.: 7 :
? SEQUENCE CHARACTERISTICS:
? LENGTH: 215 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? ORIGINAL SOURCE:
? INDIVIDUAL ISOLATE: h-spry2 cDNA
? FEATURE:
? NAME/KEY: Coding Sequence
? LOCATION: 391...1335
? OTHER INFORMATION:
?
? US-08-365-903b-7

```

Query Match	10.8%	Score 180.2	DB 3	Length 2135
Best Local Similarity	6.7%	Pred. No. 1.6e-36		
Matches 213	Conservative	0	Mismatches 133	Indels 3
				Gaps 1
Qy 661	gaagctgacaagcaactctctctgtctgtcgtcagagcgctgtggaagtgtaaatgacaagagatgt	720		
Db 898	gatttggggccttgacacgccctacagagtgtagagcattgtgcgaagtcaaatgtaagagatgc	957		
Qy 721	gaatccccccggagcgtctgtctctccgtctgcgtcgggtctgcgaacacggaggtgcctgtctaacgc	780		
Db 958	accacccaagggccttgcattcagactgacatgcatctgcacgaacgacgtctcttgcctgcggcc	1017		
Qy 781	cagagctctgtctaacatctgacgtgcagctgtctgtctgtgtgcagggacatctcttaacaatgc	840		
Db 1018	cagaaagtgattgacatgataggacattgtgtatgctgtgtgtgaaagtcctcttcatnactgt	1077		
Qy 841	acgaatgagagcaatgaagagctctctgcgcgtcacaacccctgtctctgtctccgcgtccaac	900		
Db 1078	tctnmt--gattatataagacaacatgctctgacaaacccatgcttcttgcacccagttctac	1134		
Qy 901	tgctgcgcgcgcgtgcgtctctcaatgggtgtctctccgcgtgcgtgcgtgcgtgcgtgcgtgc	960		
Db 1135	tgtttgtacagatagctcaccacacatgggtgtcatgttcccttcttggcttcttttatgtgtgt	1194		
Qy 961	taactgcgtccacacgcgtctgcgtgaagcttgagcccaagcgtgtgctacgaacgctctgcgcgc	1020		
Db 1195	tacccttcacacccaagggtgtgctttaaatttgaccacagggggtttatgacccgggttaacag	1254		
Qy 1021	ctctgtgttcgcgtctgcacagacacagaaagcgtctaatctgcgaagaagacga	1069		
Db 1255	ctctggttccgcgtctgtaaaactcaaacacagtttgctgcacaaagtctcca	1303		

RESULT 3
 US-08-965-9038-5
 Sequence 5, Application US/08965903B
 Patent No. 6060275
 GENERAL INFORMATION:
 APPLICANT: Hachcoed, Nit
 APPLICANT: Krasnow, Mark A.
 TITLE OF INVENTION: SPURUTY PROTEIN AND CODING
 TITLE OF INVENTION: SPURUTY PROTEIN AND CODING
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STEETER, 350 Cambridge Ave., Suite 250
 City: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: PASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/965,903B

RESULT 64
 US-08-965-903B-12
 : Sequence 12, Application US/08965903B
 : Patent No. 6060275
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Hacheco, Nit
 :
 : APPLICANT: Krasnov, Mark A.
 :
 : TITLE OF INVENTION: SPROUT PROTEIN AND CODING
 :
 : TITLE OF INVENTION: SEQUENCE
 :
 : NUMBER OF SEQUENCES: 20
 :
 : CORRESPONDENCE ADDRESS:
 :

[illegible]

```

RESULT 5
US-08-965-903B-1
; Sequence 1, Application US/08965903B
; Patent No. 6060275
;
; GENERAL INFORMATION:
; APPLICANT: Hachon, Nir
; APPLICANT: Krasnov, Mark A.
; TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
; TITLE OF INVENTION: SEQUENCE
; NUMBER OF SEQUENCES: 20
;

```



```

Db 903 CAGATACCATCTACACGACGCCACCGGCGCAGTACTGAGAGCGGCCAGCAT 962
Oy 167 cctagaagcctgttcttcctcagtcagtcagga--cctcagcccatggaagcccgat 224
Db 963 CATGGCCAGCCCGGTGGGCTTACAGCGGCCCGCCAGTTGACCTTCCCTGTACGGCAGCAT 1022
Oy 225 cccacagagagccccccttgaacccaactcagtcagtcagtcagccctcttgaagccg 284
Db 1023 GGGCAACGCCCGCCCGCCAGCAGCGCATCTGTGGCCACGCTGGGCCAGGGCTGTACCGCAC 1082
Oy 285 gatgtccacagcagcgtctcagacccaactcaactcactcaacttgaagtgagagac 344
Db 1083 CCTGAGCAGCACCCTGTACCGCCCGCCCTTCAACATCGCATCAACCAACAGCTGAG 1142
Oy 345 cagccatgtgagagatgaactacatagaacacctagcctgagccagccagcccaaa 404
Db 1143 CGTCTGGAGGCGACGAGTTCGCTTACGGGACACGACGACACTGCCCCAGCGCGCTGA 1202
Oy 405 gcgagaccgaggggagcccgagcgtgagcccgagcgcgcgcgtgtgagcagagatg 464
Db 1203 CCGCAGAGGCGCAGCCGTGACAGCCTGTGACGAGATCCCCCGCAGAACAAACGTGCC 1262
Oy 465 caccacattgagctctctcagcggcgcccgagcgtctgtgagcagcagcagcagca 523
Db 1263 CCCCCGCGAGGCTTACAGCAGCCGCTGAGCCAGCATGTTCGCGAGCGGCTTCA 1321

```

RESULT 12
US-08-459-504B-2
Sequence 2, Application US/08459504B
Patent No. 6075185

GENERAL INFORMATION:

```

APPLICANT: Kozziel, Michael G.
APPLICANT: Desai, Nalin M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Ewola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlio, Ellis J.
APPLICANT: Laurie, Karen L.
APPLICANT: Rohsteln, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESS: No. 6075185artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,504B
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/459,595
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3468 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..3468
OTHER INFORMATION: /product= "Full-length pure maize
OTHER INFORMATION: optimized synthetic Bt"
OTHER INFORMATION: /note= "Disclosed in Figure 3 as syn1.mze"
US-08-459-504B-2

```

Query Match 3.1%, Score 51.4; DB 3; Length 3468;
Best Local Similarity 46.8%; Pred. No. 0.0011;
Matches 196; Conservative 0; Mismatches 221; Indels 2; Gaps 1;

```

Oy 107 cggcttcagatttaccagagcgtgagcagtcgttgaacccctgacgtctcacaagc 166
Db 903 CAGATACCATCTACACGACGCCACCGGCGCAGTACTGAGAGCGGCCAGCAT 962
Oy 167 cctagaagcctgttcttcctcagtcagtcagga--cctcagcccatggaagcccgat 224
Db 963 CATGGCCAGCCCGGTGGGCTTACAGCGGCCCGCCAGTTGACCTTCCCTGTACGGCAGCAT 1022
Oy 225 cccacagagagccccccttgaacccaactcagtcagtcagtcagcccgagcgcgcgcgtgtgagcagc 284
Db 1023 GGGCAACGCCCGCCCGCCAGCAGCGCATCTGTGGCCACGCTGGGCCAGGGCTGTACCGCAC 1082
Oy 285 gatgtccacagcagcgtctcagacccaactcaactcaactcactcaacttgaagtgagagc 344
Db 1083 CCTGAGCAGCACCCTGTACCGCCCGCCCTTCAACATCGGCATCAACCAACAGCTGAG 1142
Oy 345 cagccatgtgagagatgaactacatagaacacctagcctgagccctgacccaagc 404
Db 1143 CGTCTGGAGGCGACGAGTTCGCTTACGGGACACGACGACACTGCCCCAGCGCGCTGA 1202
Oy 405 gcgagaccgaggggagcccgagcgtgagcccgagcgcgcgcgtgtgagcagagatg 464
Db 1203 CCGCAGAGGCGCAGCCGTGACAGCCTGTGACGAGATCCCCCGCAGAACAAACGTGCC 1262
Oy 465 caccacattgagctctctcagcggcgcccgagcgtctgtgagcagcagcagcagca 523
Db 1263 CCCCCGCGAGGCTTACAGCAGCCGCTGAGCCAGCATGTTCGCGAGCGGCTTCA 1321

```

RESULT 13
US-08-459-444-2
Sequence 2, Application US/08459444A
Patent No. 6121014

GENERAL INFORMATION:

```

APPLICANT: Kozziel, Michael G.
APPLICANT: Desai, Nalin M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Ewola, Stephen V.
APPLICANT: Crossland, Lyle D.

```



```

Db 963 CATGGCCAGCCCGTGGGCTTCAGCGCCCGCCAGTTCACTTCCCTGTAACGGCACCAT 1022
Oy 225 cccacagagcgcccttgatcccaactcagatcgtgctcagccctcttgacagcg 284
Db 1023 GGGCAACGCGCCCGCCAGAGGAGTCTGGCCCAAGCTGGGGCCAGGGCGGTACCGCA 1082
Oy 285 gatgccacagcgctcagagaccccaactacactcctactacatctgacaggtgaagac 344
Db 1083 CCTAGACAGCAACCTGTACCGCCCGCCCTTCAACATCGGATCAACAACAGCAAGTGA 1142
Oy 345 cagccatgtgagaaatgactacatagacaacccctagccctgagccctgacacggcccaaa 404
Db 1143 CGTCTGTAGACGCGACGAGTTCGCTTACGCGCACAGACAGCAACCTGCCAGCGCGTGA 1202
Oy 405 gcgagaccggggggggcccgagagctgggcccggacggcgccggcctgtgacagagatgt 464
Db 1203 CCGCAAGAGCGGCGACCGTGGACAGCCTGAGACAGATCCCCCCCCAGAACACAACTGCC 1262
Oy 465 caccaccattgattcctcagcgggcgcccgagctctgtgagcagcagcagca 523
Db 1263 CCCCCGCGAGGCTTCAGCCACCGCCCTGAGCCACGTGATGTTCCGACGCGCTTCA 1321

```

RESULT 15 US-09-547-422-2

Sequence 2, Application US/09547422
Patent No. 6320100

GENERAL INFORMATION:
APPLICANT: Kozziel, Michael G.

Desai, Nallini M.
Lewis, Kelly S.
Kramer, Vance C.
Warren, Gregory W.
Evola, Stephen V.
Crossland, Lyle D.
Wright, Martha S.
Merlin, Ellis J.
Launis, Karen L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS:
ADDRESSER: No. 6320100artis Agribusiness Biotechnology Research, Inc.

STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC

COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/547,422
FILING DATE: 11-Apr-2000

CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/459,595
FILING DATE: 02-JUN-1995

APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992

APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-18805H

TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587

TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 2:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 3468 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..3468
; OTHER INFORMATION: /product= "Full-length pure maize
; optimized synthetic Bt"
; /note= "Disclosed in Figure 3 as syn1T.mze"
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-547-422-2

```

```

Query Match 3.1%, Score 51.4; DB: 4; Length 3468;
Best Local Similarity 46.8%; Pred. No. 0.0011;
Matches 196; Conservative 0; Mismatches 221; Indels 2; Gaps 1;

Oy 107 cggctccagattacacagagctgagggagctgtgtgacccctgagcctcctcaagac 166
Db 903 CAGCATCACCATTCTACACGAGCCGACCGCGGAGTACTACTGAGCGCCACACAGAT 962
Oy 167 ccttagaagcctgttcttcgtaagtcagga--cctcagcccatgtgagcccgat 224
Db 963 CATGGCCAGCCCGCTGGGCTTCAGCGCCCGCCAGTTCACTTCCCTGTAACGGCACCAT 1022
Oy 225 cccacagagcgcccttgatcccaactcagatcgtatgtgtcagccctcttgacagcg 284
Db 1023 GGGCAACGCGCCCGCCAGAGCGATGCTGGCCAGCTGGGCGAGGGGTGTAACCGCAC 1082
Oy 285 gatgtcccaacagcgctcagacacccactacactcctactcattgacagagtgaaagc 344
Db 1083 CGTGACACAGCACCCTGTACCGCCCGCCCTTCAACATCGGATGAGCAACACCTGAGT 1142
Oy 345 cagccatgtgagaaatgactacatagacaacccctagccttgaccctgacacggcccaaa 404
Db 1143 CGTCTGTAGACGCGACGAGTTCGCTTACGCGCACAGACAGCAACCTGCCAGCGCGTGA 1202
Oy 405 gcgagaccggggggggcccgagagctgggcccggacggcgccggcctgtgacagagatgt 464
Db 1203 CCGCAAGAGCGGCGACCGTGGACAGCCTGAGACAGATCCCCCCCCAGAACACAACTGCC 1262
Oy 465 caccacacatgtatcctcctcagcgggcgcccgagctctgtgagcagcagcagca 523
Db 1263 CCCCCGCGAGGCTTCAGCCACCGCCCTGAGCCACGTGATGTTCCGACGCGCTTCA 1321

```

Search completed: May 11, 2002, 04:58:37
Job time: 13917 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2002, 02:16:20 ; Search time 200.42 Seconds
(without alignments)
14237.660 Million cell updates/sec

Title: US-10-082-902-1

Perfect score: 1 agtagcagcagctgagctgac.....tgccttgagggaagtgtaga 1662

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: N_Geneseq_032802.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001.DAT.*
24: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307.6	18.5	357	21	AAAC07730
2	222	13.4	369	22	AAAC07730
3	189	11.4	2138	21	AAAI4073
4	189	11.4	2195	21	AAAI4072
5	189	11.4	2330	22	AAAF3646
6	189	11.4	2393	22	AAAF3646
7	189	11.4	2553	21	AAAF3646
8	180.2	10.8	1508	21	AAAF3646
9	180.2	10.8	1508	21	AAAF3646

10	180.2	10.8	2135	19	AAV23399	Human sapiens spry2
11	179.6	10.8	2199	22	AAI59630	Human polynucleot
12	179.4	10.8	1771	21	AAAI4075	Human SPORU-1 pa
13	162.6	9.8	1603	19	AAV32400	Human sapiens spry1
14	159.6	9.6	1622	22	ABA09622	Human bone marrow
15	156.8	9.4	723	22	ABA09706	Human bone marrow
16	114	6.9	582	22	ABA59853	Human foetal liver
17	114	6.9	582	22	ABA28317	Human foetal liver
18	114	6.9	582	22	ABA08124	Human brain expres
19	114	6.9	582	22	AAK34002	Human bone marrow
20	114	6.9	582	22	AAI39725	Human bone marrow
21	98.8	5.9	3254	19	AAV32402	Probe #641 used t
22	97.4	5.9	1057	19	AACT9934	Drosophila melanog
23	97.2	5.8	1057	23	ABLO3707	Human secreted pro
24	97.2	5.8	4122	23	ABLO3706	Drosophila melanog
25	96.8	5.8	286	22	ABA23394	Human foetal liver
26	96.8	5.8	286	22	ABA38196	Probe #16662 for g
27	96.8	5.8	286	22	AAK20818	Human brain expres
28	96.8	5.8	286	22	AAK46970	Human bone marrow
29	96.8	5.8	286	22	AAI52806	Probe #21497 used
30	91	5.5	601	22	AAH81506	Human differential
31	79.4	4.8	760	21	AAAI4071	Human SPORU-1 pa
32	79.4	4.8	2030	21	AAAI4071	Human SPORU-1 pa
33	64.8	3.9	300	19	AAV32401	Human sapiens spry3
34	64.4	3.9	255	22	AAH82143	Rat differential t
35	51.4	3.1	3474	14	AAQ39018	Sequence of pure m
36	51.4	3.1	3474	18	AAQ39018	Maize optimised-B
37	51.4	3.1	3474	19	AAV6191	Maize optimised DN
38	50.2	3.0	795	19	AAV5830	FLGA insert stabl
39	47.2	2.8	5857	21	AAAS8471	Nucleotide sequenc
40	46.2	2.8	493	21	AAZ50580	Human epidermal pr
41	45	2.7	1337	20	AAZ17263	Human gene express
42	44.6	2.7	1459	21	AAA02528	Human colon cancer
43	44.6	2.7	6652	23	ABLI3170	Drosophila melanog
44	44.4	2.7	1879	12	AAO11126	Sequence encoding
45	44.2	2.7	10732	21	AAAI0594	Gene encoding a su

ALIGNMENTS

RESULT 1	AAAC07730	standard; cDNA; 357 BP.
ID	AAAC07730	standard; cDNA; 357 BP.
AC	AAAC07730:	
DT	06-OCT-2000	(first entry)
DE	Human secreted protein 5' EST, SEQ ID NO: 11805.	
XX	Human: 5' EST: expressed sequence tag; secreted protein; cDNA isolation;	
XX	gene therapy; chromosome mapping; ss.	
OS	Homo sapiens.	
XX	EP1033401-A2.	
PN	06-SEP-2000.	
PD	21-FEB-2000; 2000EP-0200610.	
XX	26-FEB-1999; 9905-0122487.	
XX	(GENSET) GENSET.	
PA	Dumas Mline Edwards J, Duclert A, Giordano J;	
PI	WPI; 2000-500381/45.	
XX	New nucleic acid that is a 5' expressed sequence tag (5' EST) for	
PT	obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for	
PT	diagnostic, forensic, gene therapy and chromosome mapping procedures -	

XX Claim 1; SEQ ID 11805; 71bp + CD-ROM; English.
PS
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or POLYA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 3' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
XX Sequence 357 BP; 69 A; 124 C; 98 G; 65 T; 1 other;

Query Match	18.58;	Score 307.6;	DB 21;	Length 357;
Best Local Similarity	99.78;	Pred. No. 3.2e-67;		
Matches 307; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

	RESULT
XX	2
XX	AAS27967/c
XX	ID
XX	AAS27967 standard; cDNA; 369 BP.
XX	AC
XX	AAS27967;
XX	DT
XX	07-NOV-2001 (first entry)
XX	DE
XX	Novel cDNA encoding for human respiratory antigen #99.
XX	XX
XX	Human; respiratory antigen; respiratory disorder; throat disorder;
XX	lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
XX	anti allergic; anti asthmatic; anti inflammatory; olfactory;
XX	respiratory active; ss.
XX	OS
XX	Homo sapiens.
XX	PN
XX	WO20015448-A1.
XX	XX
XX	02-AUG-2001.
XX	PD
XX	17-JAN-2001; 2001MO-US01333.
XX	PF
XX	XX

[illegible]

PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0246174.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250300.
 PR 01-DEC-2000; 2000US-0250360.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM,
 DR MPI: 2001-476224/51.
 DR P-PSDB: AAU17783.
 PT Isolated polypeptide for treating, preventing and/or prognosing
 disorders related to the respiratory system including respiratory

PT cancers and also for testing and detection e.g. diagnosis -
 XX
 XX Claim 4; SED ID No 109; 546pp; English.
 PS
 XX The present invention relates to the isolation of novel human
 CC respiratory antigens (AAU17685-AAU17975), and cDNA and genomic
 CC sequences encoding for these polypeptides. The sequences of the
 CC invention are useful for preventing, treating and/or prognosing
 CC disorders related to the respiratory system including throat
 CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),
 CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,
 CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of
 CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences
 CC of the invention are useful in gene therapy and antisense therapy.
 CC AA527869-AA528159 encode for novel human respiratory antigens.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 369 BP; 91 A; 85 C; 99 G; 90 T; 4 other:
 Query Match 13.4%; Score 222; DB 22; Length 369;
 Best Local Similarity 97.8%; Pred. No. 8.3e-46;
 Matches 225; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1226 cccactccctaccctccagcttgcgaataacagagaccacacagctacccgtattcc 1285
 Db 235 cccactccctaccctccagcttgcgaataacagagaccacacagctacccgtattcc 176
 QY 1286 ccaagatgataaagaacacatttggtgttttttcagaggtcctgaacttggtcaaa 1345
 Db 175 CCAAGATGATGAAGAAGACACTTGGGCTTTTTCAGGGCTCGAATCTTGCTCAAA 116
 QY 1346 cacagaaatgcagagggcaggtgtgtttggggggaatttttttttcagaagacagaa 1405
 Db 115 CAGACAAATGACGGGCGACGCTGTGTTGGGGGCAAAATTTTCTTTTCAGAAAGACAGA 56
 QY 1406 cacagatgtgacacatacgcgaacatgcagctgtcgtgaatgccttccc 1455
 Db 55 AACAGATGTGACACATATCCGGAACATGCAGCTGCTGAATCTGTGCC 6
 RESULT 3
 ID AA14073 standard; cDNA; 2138 BP.
 XX
 AC AA14073;
 XX
 DT 21-JUL-2000 (first entry)
 XX
 DE Human SPROUTY-1 partial cDNA, SEQ ID NO:6.
 XX
 KW SPROUTY-1; spry1; fibroblast growth factor; RGF inhibitor;
 KW signalling; tumorigenesis; neurogenesis; metapneurogenesis; cancer;
 KW tumour; prostate; adrenal gland; gene therapy; ss.
 OS Homo sapiens.
 XX
 FH Key 2.889 Location/Qualifiers
 FT CDS /tag a
 FT /partial
 FT /product= "Human SPROUTY-1 protein fragment, SEQ ID NO:5
 XX (AAV87593)"
 XX /note= "No initiation codon given in the specification"
 XX
 PN WO200015781-A1.
 XX
 PD 23-MAR-2000.
 XX
 PF 10-SEP-1999; 99MO-US20638.
 XX

PR 11-SEP-1998; 98US-0099936
PR 11-SEP-1998; 98US-0099936
PR 11-SEP-1998; 98US-0099936PA (ELIL) LILLY & CO ELI.
xx

PI Edmonds BT;

DR WPI: 2000-271422/23.

DR P-PSDB; AA87593

PT New polynucleotide encoding a 319 amino acid sequence, used as a
PT fibroblast growth factor (FGF) antagonist for treating prostate cancer
PT and inhibiting angiogenesis in mammals -

Example 1: Flg 6; 93pp; English

This sequence represents a cDNA clone encoding a human SPROUTY-1 protein (also referred to as spry1 in the specification). SPROUTY-1 proteins (AA057591-1/87595) are antagonists of members of the fibroblast growth factor (FGF) family. The SPROUTY-1 proteins of the invention, and fragments thereof may be used in mammals to limit FGF levels *in vivo* and to control FGF-mediated signalling. They may be used to counteract FGF overexpression associated with certain types of tumours, and to control the growth, development and differentiation of any cell responsive to FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to control tumorigenesis, neurogenesis, metanephrogenesis, and particularly angiogenesis. They may also be used to treat conditions associated with SPROUTY-1 underexpression including polycystic kidney disease, nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial insufficiency and various cancers, particularly prostatic tumours. They may additionally be used to treat a mammal with an adrenal gland pathology involving hypersensitive responses to FGF. Nucleic acid sequences encoding SPROUTY-1 proteins may be used to recombinantly express SPROUTY-1 proteins, as a source of primers and probes, and for antisense suppression of SPROUTY-1 gene expression.

Sequence 2138 BP; 590 A; 454 C; 433 G; 661 T; 0 other;

Query Match	11.48;	Score 189;	DB 21;	Length 2138;
-------------	--------	------------	--------	--------------

Best Local Similarity 59.48; Pred. NO. 2.8e-37;
Matched 343; Conserved 6

Matches 345; conservative 0; mismatches 225; indels 9; gaps 1,

[illegible][illegible]

RESULT

ID AA014072 standard; cDNA; 2195 BP

AC AAA14072;

DT 21-JUL-2000 (first entry)

Human SPROUTY-1 partial cDNA, SEQ ID NO:4.

kw sprouty-1; spry1; fibroblast growth factor; egf inhibitor; signal trans. tumor suppressors; notochord homeobox

cumhuri; prosclate; aurelian glandu; gene chierapy; ss.

XX
XX

FT	2. .886
CDS	

ET
/partial

ET (AY87592)*

XX
XX
XXXXXXX
XXXXXXX

XX 22-MAR-2000
PD

10-SEP-1999-

11-SEP-1998. 09115-00999936

PA (ELL) LILLY & CO ELL.

PI Edmonds BT;

DR WPI; 2000-2

XX

PT fibroblast growth

XX
XX

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

This sequence represents a cDNA clone encoding a human SPKRY-1 protein (also referred to as SPY1 in the specification). SPKRY-1 proteins (AA87591-1/87595) are antagonists of members of the fibroblast growth factor (FGF) family. The SPKRY-1 proteins of the invention, and fragments thereof may be used in mammals to limit FGF levels *in vivo* and to control FGF-mediated signalling. They may be used to counteract FGF overexpression associated with certain types of tumours, and to control the growth, development or differentiation of any cell responsive to FGF-mediated signalling. For example, SPKRY-1 proteins may be used to control tumorigenesis, neurogenesis, metanephrogenesis, and particularly angiogenesis. They may also be used to treat conditions associated with SPKRY-1 underexpression including polycystic kidney disease, nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial insufficiency and various cancers, particularly prostatic tumours. They may additionally be used to treat a mammal with an adrenal gland pathology involving hypersensitive responses to FGF. Nucleic acid sequences encoding SPKRY-1 proteins may be used to recombinantly express SPKRY-1 proteins, as a source of primers and probes, and for

Db 637 cctggttcgctgtaaaactcaaacacagttgctgcaagttccca 685

RESULT 10
AAV32399 standard; cDNA to mRNA, 2135 BP.

AC AAV32399;
XX
DT 11-SEP-1998 (first entry)
XX
DE Homo sapiens spry2 gene.
XX
KW spry2; sprouty 2 protein; angiogenesis; treatment; cancer;
KW lateral inhibitory signal; epithelial tube branching; inhibition;
KW neovascularisation; psoriasis; angioidfibromas; immune; non-immune;
KW inflammation; rheumatoid arthritis; capillary proliferation;
KW atherosclerotic plaques; haemangiomas; Kaposi's Sarcoma;
KW endometriosis; scar formation; wound healing; ds.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT CDS 391..1338
FT /*tag= a
FT /product= sprouty 2 protein

W09820032-A1.
XX
XX 14-MAY-1998.
XX
XX 07-NOV-1997; 97MO-US20398.
XX
XX
XX 07-NOV-1996; 96US-0030232.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX
XX Hachon N, Krasnow MA;
XX
XX WPI: 1998-268662/25.
XX P-PSDB; AAM46792.

New isolated human sprouty polypeptide(s) - used to develop products for treating disorders involving angiogenic dysfunction, e.g. cancers, diabetic retinopathy, glaucoma, psoriasis or inflammation

Claim 12: Page 53-55; 90pp; English.

The sequence is that encoding human sprouty 2 protein. The protein functions as a lateral inhibitory signal during epithelial tube branching. It can be used for inhibiting angiogenesis and neovascularisation in the treatment of e.g. cancers, diabetic retinopathy, retrolental fibroplasia, neovascular glaucoma, psoriasis, angiofibromas, immune and non-immune angiofibromas, immune and non-immune inflammation (including rheumatoid arthritis), capillary proliferation within atherosclerotic plaques, haemangiomas, Kaposi's Sarcoma, endometriosis and unwanted scar formation in wound healing. Products of the DNA and proteins can also be used for detection and diagnosis.

Sequence 2135 BP; 583 A; 474 C; 488 G; 590 T; 0 other;

Query Match 10.8%; Score 180.2; DB 19; Length 2135;
Best Local Similarity 66.7%; Pred. NO. 4.5e-35;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

QY 661 gagctgcagaagcactctctgctgagagcctgctggagagtgtaaatgaagagtgct 720
II III I IIII I IIII IIII IIII IIII IIII IIII IIII
DB 898 gatttggcctgcagcagctacagtgctggagactgctggcaagtgaatgaagagtgct 957
721 gcatcccccggagctgctctctctgctgctgctgcaaacaggagtgctgctgctcagcc 780

Db 958 acctaccagaagcctctgcacatcagctgcatcgtgcagaaagcagctcttctgctggcc 1017
II IIII II I IIII II IIII IIII IIII IIII IIII IIII IIII
QY 781 cagactctgttcaactatggtgacagctgcatctgtttgtggcagggcatctcttaccactgc 840
IIII II IIII IIII IIII IIII IIII IIII IIII IIII IIII
DB 1018 cagaacgttgattgactatggagactgtgtatgctgtgtgfaaaggctctcttaccactgt 1077
QY 841 acgaatgagcagcagtgaggtcctctgctgacacacccctctctctgctccgcgtccac 900
IIII II IIII IIII IIII IIII IIII IIII IIII IIII IIII
DB 1078 tctaact--gatgatggagcagcactgtgctgacacacatgttcttctgacagctctcac 1134
QY 901 tgcctgcgcgcgtgctctctctcatggtgtcctctcctcgtggtgctcctgctctgc 960
II II I IIII IIII IIII IIII IIII IIII IIII IIII
DB 1135 tgtgtacacagatggtcagccatggtgtgtcatgtccctcttcttctgtttatggtgt 1194
QY 961 tactgtcctgcacacggctgctgctgtaagctgagcagcgtggtgctacagcgtctgcgcgc 1020
IIII II IIII IIII IIII IIII IIII IIII IIII IIII
DB 1195 tactctccagcacaagggttgccttaattgtgcaggggtgttatgaccgggttaacagg 1254
QY 1021 cctggttgcgcgtgcagacagcagcaagcagcgtcatctgcaagcagcca 1069
IIII IIII IIII IIII IIII IIII IIII IIII IIII
DB 1255 cctggttgcgcgtgtaaaactcaaacacagtttgcgtgcaagttccca 1303

RESULT 11
AAI59630/C
ID AAI59630 standard; cDNA; 2199 BP.
XX
XX AAI59630;
XX
XX
XX 22-OCT-2001 (first entry)
XX
XX
XX Human polynucleotide SEQ ID NO 3619.
XX
XX
XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
XX
XX Homo sapiens.
XX
XX
XX W0200153312-A1.
XX
XX
XX 26-JUL-2001.
XX
XX
XX 26-DEC-2000; 2000MO-US34263.
XX
XX
XX 21-JAN-2000; 2000US-0488725.
XX
XX 25-APR-2000; 2000US-0552317.
XX
XX 09-JUL-2000; 2000US-0598042.
XX
XX 19-JUL-2000; 2000US-0620312.
XX
XX 03-AUG-2000; 2000US-0653450.
XX
XX 14-SEP-2000; 2000US-0662191.
XX
XX 19-OCT-2000; 2000US-0693036.
XX
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEO INC.
XX
XX
XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Driamanac RT;
XX
XX WPI: 2001-442253/47.
XX P-PSDB; AAM40474.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 3619; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and

Db 182 gagagcatggtggaataatggaacctgcatgtgcttagtcaaggagcatcttctacacatc 241
 Oy 841 acgcatgagcagatgagcgtctctgctgacacacgtctctctccgctccac 900
 Db 242 tccatgacgagcagaggaattccattcaagtaactctgctctcttcaacaacac 301
 Oy 901 tgcgtgacgacgctgctctctcaatgagtgctctctccgtggtgctgctgctcgc 960
 Db 302 tgcgtgctctgatacctgtgtatggaagcagtgctcttatttacctgtctactctgt 361
 Oy 961 taccgtgctgacacgctgctgctgagcgtgacgagcgtgctgctgacacgtctgacgc 1020
 Db 362 tatctctctgcttaagatgagcctggaagcgtgctgaggaagtgttatgactggatccatcgc 421
 Oy 1021 cctggttgcgcgtgctgacgaacgaacagcgtcatctgcaagaacgacgcgcg 1073
 Db 422 ccaggtgctgacagatgatagaactccacactgctctatgttaagctgagagactg 474

RESULT 13
 AAV32400
 ID AAV32400 standard; cDNA to mRNA; 1603 BP.
 AC AAV32400;
 DT 11-SEP-1998 (first entry)
 DE Homo sapiens spry1 gene.
 KW spry1, sprouty 1 protein; angiogenesis; treatment; cancer;
 KW lateral inhibitory signal; epithelial tube branching; inhibition;
 KW neovascularisation; diabetic retinopathy; retrolental fibroplasia;
 KW neovascular glaucoma; psoriasis; angiodiomas; immune; non-immune;
 KW inflammation; rheumatoid arthritis; capillary proliferation;
 KW atherosclerotic plaques; haemangiomas; Kaposi's Sarcoma;
 KW endometriosis; scar formation; wound healing; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..420
 FT /tag= a
 FT /product= sprouty 1 protein
 PN W09820032-A1.
 PD 14-MAY-1998.
 PF 07-NOV-1997; 97MO-US20398.
 PR 07-NOV-1996; 96US-0030232.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Hachon N, Krasnow MA;
 DR WPI; 1998-286862/25.
 DR P-PSDB; AAM48793.
 PT New isolated human sprouty polypeptide(s) - used to develop products
 PT for treating disorders involving angiogenic dysfunction, e.g.
 PT cancers, diabetic retinopathy, glaucoma, psoriasis or inflammation
 PS Claim 14; Page 50-51; 90pp; English.

The sequence is that encoding human sprouty 1 protein. The protein
 functions as a lateral inhibitory signal during epithelial tube
 branching. It can be used for inhibiting angiogenesis and
 CC neovascularisation in the treatment of e.g. cancers, diabetic
 CC retinopathy, retrolental fibroplasia, neovascular glaucoma, psoriasis,
 CC angiodiomas, immune and non-immune inflammation (including rheumatoid
 CC arthritis), capillary proliferation within atherosclerotic plaques,
 CC haemangiomas, Kaposi's Sarcoma, endometriosis and unwanted scar

CC formation in wound healing. Products of the DNA and proteins
 CC can also be used for detection and diagnosis.
 XX
 SQ Sequence 1603 BP; 411 A; 320 C; 316 G; 556 T; 0 other;

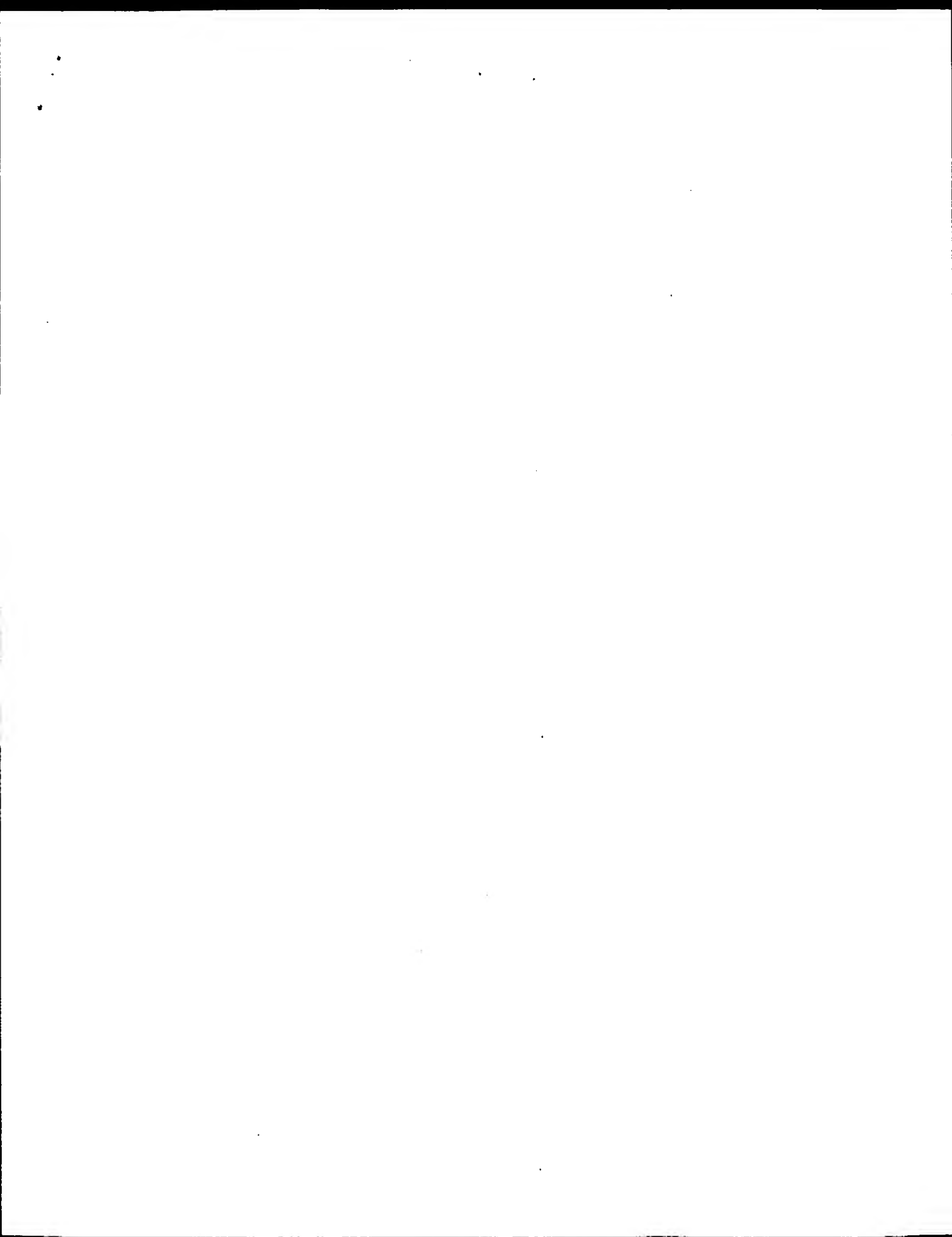
Query Match 9.8%; Score 162.6; DB 19; Length 1603;
 Best Local Similarity 65.3%; Pred. No. 1e-30;
 Matches 254; Conservative 0; Mismatches 134; Indels 1; Gaps 1;

Oy 685 tgcgagcgcgtgagagatgtaataatgcaagagtgatgacccccggagctgctcc 744
 Db 1 tgcgaacagtgtggaagtgcgaagtgtggaatgacactgctccagagacctaccacc 60
 Oy 745 tgcgtgagctgcaacacagagtgctgtgtcagcccaagactgtgtcaactatgacag 804
 Db 61 tgttgccgttaacccgacgtgctctgtcctgtgagagcagatgtgatatggaacc 120
 Oy 805 tgcagtgtgtgtgtagagggatctcttaccatgcaacagatgagagcgtgagggctcc 864
 Db 121 tgcagtgtgtgttagaagggacatctcttaccacacgtcccaatgagcagagggattcc 180
 Oy 865 tgcgtgacacacccctgctctgctccgcgtccacactgtgctgcccgtgttctcatg 924
 Db 181 tattcagataatcctctgctccgttccacatccacactgtgctctagataccgtgtatg 240
 Oy 925 ggtgctctctcggtgtgtgtgctgctgctgctgctgctgctgctgctgctgctgctg 984
 Db 241 ggaagcagatcttatttatttacctgtgctgctgctgctgctgctgctgctgctgctg 300
 Oy 985 aagctgcccagcgtgtgctacagccgtctgcccgtgtgtgctgctgcaagacag 1044
 Db 301 aagct-gtcagaggtgttatgtacgtgacatgcagccaggtgtcagatgttaagaatcc 359
 Oy 1045 aacagcgtcatctgcaaacgacgcgcg 1073
 Db 360 aacacgtctatgttaagctgagagctg 388

RESULT 14
 ABA09622
 ID ABA09622 standard; DNA; 1622 BP.
 AC ABA09622;
 DT 15-JAN-2002 (first entry)
 DE Human bone marrow expressed oligonucleotide SEQ ID NO: 48.
 KW Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnery;
 KW antiinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
 KW antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
 KW antidiacer; fungicide; antidiabetic; antiaesthetic; antiallergic;
 KW immunostimulant; analgesic; cerebroprotective; antianemic; infection;
 KW nervous system disorder; autoimmune disorder; inflammation; allergy; ds.
 OS Homo sapiens.
 FH W0200174836-A1.
 FT 11-OCT-2001.
 PD 30-MAR-2001; 2001MO-US10472.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PR 23-AUG-2000; 2000US-0649267.
 PR 30-NOV-2000; 2000US-250583P.
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Dzumanac RT, Ford JE, Boyle BJ;

Db 329 atgataagacaactgtgctgaatgagccctgctcttgggacctagtcttctgtcc 388
Oy 911 gctgtccctcatatgggtgctctccgtggatgcccctgacctgtctgtactgtcgtg 970
||| | ||| | ||| | | ||||| ||||| |||||
Db 389 gctgggcaagcctatgagcctcatctccctcttctactacctgtgctgtactgtccta 448
Oy 971 ccacggctgctgtaagctggcccagcgttgctaaagacgltctgcccgtccctgtgtgcc 1030
|| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
449 ccgctggtatgctccatctgtggcaacagggctatgataagcctccggcgaccaggtgcc 508
Oy 1031 gctgcaagcacagacagcgtcatctgtcaaa 1062
||| ||| | | | | | ||| |
Db 509 gctgcaagagggacacacacactgtgtgcaga 540

Search completed: May 11, 2002, 05:34:20
Job time: 11880 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2002, 01:05:50 ; Search time 1753.28 Seconds
(without alignments)
19837.050 Million cell updates/sec

Title: US-10-082-902-1

Perfect score: 1662
1 agtagcagagctgagctgac.....tgctctggggaaggtgtaga 1662

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenBank:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pt:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: gb_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1591.2	95.7	4934	9	AF227516
2	888	53.4	1405	10	AF176906
3	869.8	52.3	1131	10	AB019280
4	827.8	49.8	155614	2	AC012620
5	822.4	49.5	183042	2	AC099211
6	734.2	44.2	7053	2	AF227517
7	291.8	17.6	2890	5	AF371368
8	205	12.3	2502	10	AF176903
9	201.8	12.1	176250	2	AF167086
10	194.6	11.7	1889	5	AF176904
11	189	11.4	2330	6	AX058191
12	189	11.4	2520	6	AC026960
13	189	11.4	166451	2	AC025511
14	189	11.4	176959	2	AC026402
15	180.2	10.8	1508	6	AX035193
16	180.2	10.8	1508	6	AX342047
17	180.2	10.8	1508	6	AX342107
18	180.2	10.8	1918	9	BC015745
19	180.2	10.8	2135	9	AF039843
20	180.2	10.8	141175	2	AC012110
21	180.2	10.8	179470	2	AL162211
22	180.2	10.8	191652	9	AL354668
23	173	10.4	1604	9	AF041037
24	169	10.2	948	10	AF153084
25	167.4	10.1	1519	10	AF176905
26	164	9.9	155614	2	AC012620
27	161.2	9.7	168066	2	AC025226
28	159.6	9.6	135484	2	AL356301
29	159.6	9.6	196456	2	AC009620
30	159.6	9.6	240000	9	HSN271735
31	155.6	9.4	945	5	AF331824
32	155	9.3	924	5	AF331825
33	98.8	5.9	3254	3	AF039842
34	98	5.9	73889	2	AC009983
35	97.2	5.8	91318	2	AC015391
36	97.2	5.8	113816	2	AC004358
37	97.2	5.8	172271	3	AC010018
38	97.2	5.8	178018	3	AC091206
39	91	5.5	601	6	AX210373
40	83.8	5.0	1465	5	AF177875
41	69	4.2	125020	9	AF429315
42	64.8	3.9	300	9	AF041038
43	64.4	3.9	255	6	AX211010
44	63.6	3.8	125020	9	AF429315
45	62	3.7	46448	9	AC005789

ALIGNMENTS

RESULT 1

LOCUS AF227516 4934 bp mRNA linear PRI 02-FEB-2001

DEFINITION Homo sapiens sproutly-4A mRNA, complete cds.

ACCESSION AF227516

VERSION AF227516.1 GI:12655912

KEYWORDS

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 4934)

Leeksa,O.C., van Achterberg,T.A.E., Spaargaren,M., Kr. von dem

Borne,A.E.G., Pannekoek,H. and de Vries,C.J.M.

Identification of a novel human Sproutly homologue

Unpublished

2 (bases 1 to 4934)

de Vries,C.J.M., Leeksa,O.C., van Achterberg,T.A.E. and

Pannekoek,H.

Direct Submission

Submitted (21-JUN-2000) Biochemistry, Academic Medical Center,

Weibergdreef 15, Amsterdam 1105 AZ, The Netherlands

Location/Qualifiers

[illegible][illegible]

*	569	1350: contig of 782 bp in length	*		
*		gap of unknown length	*	30200	31063: contig of 864 bp in length
*	1351	1578: contig of 228 bp in length	*		gap of unknown length
*		gap of unknown length	*	31064	32094: contig of 1031 bp in length
*	1579	2731: contig of 1153 bp in length	*		gap of unknown length
*		gap of unknown length	*	32095	32924: contig of 830 bp in length
*	2732	3366: contig of 635 bp in length	*		gap of unknown length
*		gap of unknown length	*	32925	33506: contig of 582 bp in length
*	3367	4347: contig of 981 bp in length	*		gap of unknown length
*		gap of unknown length	*	33507	34428: contig of 922 bp in length
*	4348	5253: contig of 906 bp in length	*		gap of unknown length
*		gap of unknown length	*	34429	35484: contig of 1055 bp in length

QY	186	cytaacag tccagagaccctccagagccca ttgagcccccat tcccaaaagccccctttag	245
Db	181	cgtacgtatccagagaccttccagcccccatgtatgaccccccatccacagagccccctttttagc	240
QY	246	ttcccaactcag tca tgg tccagccccctctctcttgacagccgagtgtcccaagccggtccca	305
Db	241	ttcccaactcag tca tgg tccagccccctctctcttgacagccgagtgtcccaagccggtccca	300
QY	306	gcaacccatccacaa tccaaacca ttgaacaag tgaagaacagccaatgttgaaatgacta	365
Db	301	gcacccactcacatccttaccctatttaccagstgaagaccgcatgttgaaatgacta	360
QY	366	catatagaacacctatgcttgcttgcttgacacagcgagcccaagcgagccccgggagcccc	425
Db	361	catatagaacacctatgcttgcttgcttgacacagcgagcccaagcgagccccgggagcccc	420
QY	426	agaagctggcccccgacgcccccgcccgctgtgtacacagaatgttcaaccaaatgttatctctt	485
Db	421	agaagctggcccccgacgcccccgcccgctgtgtacacagaatgttcaaccaaatgttatctctt	480
QY	486	cagagcgagggccccccagcctctgttgagcagacgacgacatctcttgacccaagcgctctt	545
Db	481	cagagcgagggccccccagcctctgttgagcagacgacgacatctcttgacccaagcgctctt	540
QY	546	agacacaca tgggacccaaccaaccg tggcttgacacaggtctcaaccaaggtctgtgcacatca	605
Db	492	-----	491
QY	606	ggcccaaggtgg tccacatcgcaagcgcttgagactcaagggccccggggtcccaaccggagct	665
Db	492	-----	491
QY	666	ggacacagcaactcttgctgttgagagcctgtgaggaaatgaatlgcaaggaatgtgcatc	725
Db	492	-----	491
QY	726	ccccccgagtgctgctctccctgcgtgggtctcgcgaaccaaggagtgctctgttccaagccccagac	785
Db	492	-----	491
QY	-786	tctgtgtaacata tggacagctgacatgtgtgttggttgacaggtactctctacacactlgacaa	845
Db	492	-----	491
QY	846	ttagagacgtatgagggctccctcggtctgacacccccctgcctcctgcctcccgctccaactgtgt	905
Db	492	-----	491
QY	906	cagccgctgtgctcttca tgggtgtctctcgttggtgtgctgccccctgtcctgtactact	965
Db	492	-----cctgctctgctgact	506
QY	966	ggctgcacacccgctctgtgttgaagttggccccagagtggtgtcaagacagctctgscgccccgg	1025
Db	507	ggctgcacacccgctctgtgttgaagttggccccagagtggtgtcaagacagctctgscgccccgg	566
QY	1026	ttgcgcgtctgacaagacacgaaacagcg tca tctgcgaagacagccagcggtgagtatgccaagac	1085
Db	567	tttccctctgcacaaacacacgaaacagcg tca tctgcgaagacagccagcggtgagtatgccaagac	626
QY	1086	caaacagggcccccaagaagcctttctgaagaattgtgtcgaagccccagtgctcgtcccttgaa	1145
Db	627	caaacagggcccccaagaagcctttctgaagaattgtgtcgaagccccagtgctcgtcccttgaa	686
QY	1146	acctgtgtctctcttgacatctaaagaactgcaatcaaggtlcaagaagtttttttagcctctgt	1205
Db	687	acctgtgtctctcttgacatctaaagaactgcaatcaaggtlcaagaagtttttttagcctctgt	746
QY	1206	aggtctgaactgtgtagtctgtcccaacccccctaccccccaattcggaaaaataacgaagacac	1265
Db	747	aggtctgaactgtgtagtctgtcccaacccccctaccccccaattcggaaaaataacgaagacac	806
QY	1266	cacacagtaacctgtatctccccaagaatgataagaagaacactlvtggggcttttttttagagg	1325

D	b		807	CACCAGTACCCGTATTTCCCCAAGCTATTATAAACACTTTGGGCGTTTTTCACGGC	866
OY	1326	tccgtaaacttggtgcaacaacagacaatgcaggggcaggglvtggttctttgggggaattt	1385		
D	867	TCTGAAACAATTCTTGCAAAACAGAACAAATGCAGGGGGCAGGGTGTGGTTGGGGGAATTT	926		
OY	1386	tctcttttcagaagaacacagatgtlgacacataacccgaaaactgcagtctcta	1445		
D	927	TTCCTTTTCAGAAGACAGAACACATATGTGGACACATATCCGGAACACTGCAGCTGCTTCA	986		
OY	1446	atgccttcccagccccctctcttccctccctccctccgccgccccccctctcttcat	1505		
D	987	ATGCCCTCCCAGGCCCTCTCTTCTCCTCCGTCCTCCGCCGCCGCCCTCTCTCTTTCAT	1046		
OY	1506	tgtcttgtaacctcacagagactgctgcgcgggaggaaatlgtlaactagtagtacagagt	1565		
D	1047	TGCTTTGGCGTCTCACAGAGACTACTGCTCCTGGAGGAAATGTTAAGTAGTACAGGGT	1106		
OY	1566	aaccttaagaagaacacctggagtgcttatcaactcttctctctcccatctaatcca	1625		
D	1107	ACCTTTAAAGAAACCCCTTGAGGTCTCTATACCTCTCTCTCTCCCTCCCATCTACTCCA	1166		
OY	1626	ccccaacttgcctgatgtlctlttgsgaagggtlaga	1662		
D	1167	CCCCACTTGTCCCTGATGTCTGGGGAAGGTGAGA	1203		
RESULT	7				
LOCUS	AF371368	2890 bp	mRNA	linear	VMT 14-SEP-2001
DEFINITION	Danio rerio Sproutly4 (sproutly4) mRNA, complete cds.				
ACCESSION	AF371368				
VERSION	AF371368.1	GI:13991912			
KEYWORDS	.	zebrafish.			
SOURCE	Danio rerio				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio. 1 (bases 1 to 2890) Furthauer,M., Reifers,F., Brand,M., Thisse,B. and Thisse,C. sproutly acts in vivo as a feedback-induced antagonist of FGF signaling in zebrafish development 128 (12), 2175-2186 (2001)				
JOURNAL	MEDLINE	21384945			
PUBMED	2	11493538			
REFERENCE	2 (bases 1 to 2890)				
AUTHORS	Fuerthauer,M., Reifers,F., Brand,M., Thisse,B. and Thisse,C.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-Apr-2001) Institut de Genetique et Biologie Moléculaire et Cellulaire, 1 rue Laurent Fries, Ilkirch, France 67000				
FEATURES					
Source	Location/Qualifiers				
	1..2890	/organism="Danio rerio"			
gene	1..2890	/db_xref="taxon:7955"			
	1..2890	/gene="sproutly4"			
CDS	623..1555	/gene="sproutly4"			
		/note="FGF signaling antagonist"			
		/codon_start=1			
	/product="Sproutly4"				
	/protein_id="AAK51555.1"				
	/db_xref="GI:13991913"				
	/translation="MESRVPHHIGVSSIMVOPLDSEVPGRLOHPITVYPIDOKK AHLENDVIDPAVISQEPSHKNPGEQVLLGAPHHNLSRCPVATTPHWISPS GPSISSSTSDORLDLHAAPTPLYDDPTNGSHSGTLAADPKLTISKNTKLTA ALPEKKRYHLICEKRCKRCHETLPRLPYSCWNCNOCLCSAQNLVDYSYCMLVK GVFIHTCEDBEDRGSCADRPCSCSHSNCCAMSFMAVAIVLYELCVLTYPATPGCAKLSQ KITDYASPRGRCKYSTOSCYAEIAIKOPEKMS"				
BASE COUNT	801 a	702 g	647 t	740 c	

QY	1211	AGTCCACCTGCTCTCTCCAAATACCTGTCATGACATGGACCCCTCTCTTTTGTGCTACCTGCT	1270
QY	953	tgctctgactactgctctgcaacgagctgcgtgaagctgagccagctgctgactacgacgctc	1012
Db	1271	TGCTCTGTCTACCTCCCGGCAAGGCTCCTCAAGCTCTGCGAGGCGCTCTTACGATCGCA	1330
QY	1013	tgagccgcccctggttcgcgctgcaagacacagaaacgagtcattctgcaagcagcagcg	1072
Db	1331	CCACAGCCCTGCTGCTGCTACAGTGTAAACACTCCAAACACTGCTCTATTGTAACTGAGACGT	1390
QY	1073	g 1073	
Db	1391	g 1391	

RESULT 9

AL627086 176250 bp DNA linear HTG 30-JAN-2002

LOCUS AL627086 Mus musculus chromosome 3 clone RP23-477P14, *** SEQUENCING IN PROGRESS ***, in unordered pieces.

ACCESSION AL627086

VERSION AL627086.13 GI:18476818

KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (sites)

AUTHORS Hammond, S.

TITLE Direct Submission

JOURNAL Submitted (23-JAN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT On Feb 1, 2002 this sequence version replaced gi:18375847.

Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquerry@sanger.ac.uk

----- Project Information -----

Center project name: bm477P14

----- Summary Statistics -----

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 175868 bases at least 040

Consensus quality: 175954 bases at least 030

Consensus quality: 175997 bases at least 020

Insert size: 176050; sum-of-contigs

Insert size: 176864; 5.7% error; agarose-fp

Quality coverage: 12.63x in Q20 bases; sum-of-contigs Quality coverage: 12.58x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

FEATURES

SOURCE

Location/Qualifiers

1..176250

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="3"

/clone="RP23-477P14"

/clone_11b="RPC1-23"

1..70998

/note="assembly_fragment:02320

fragment_chain:1

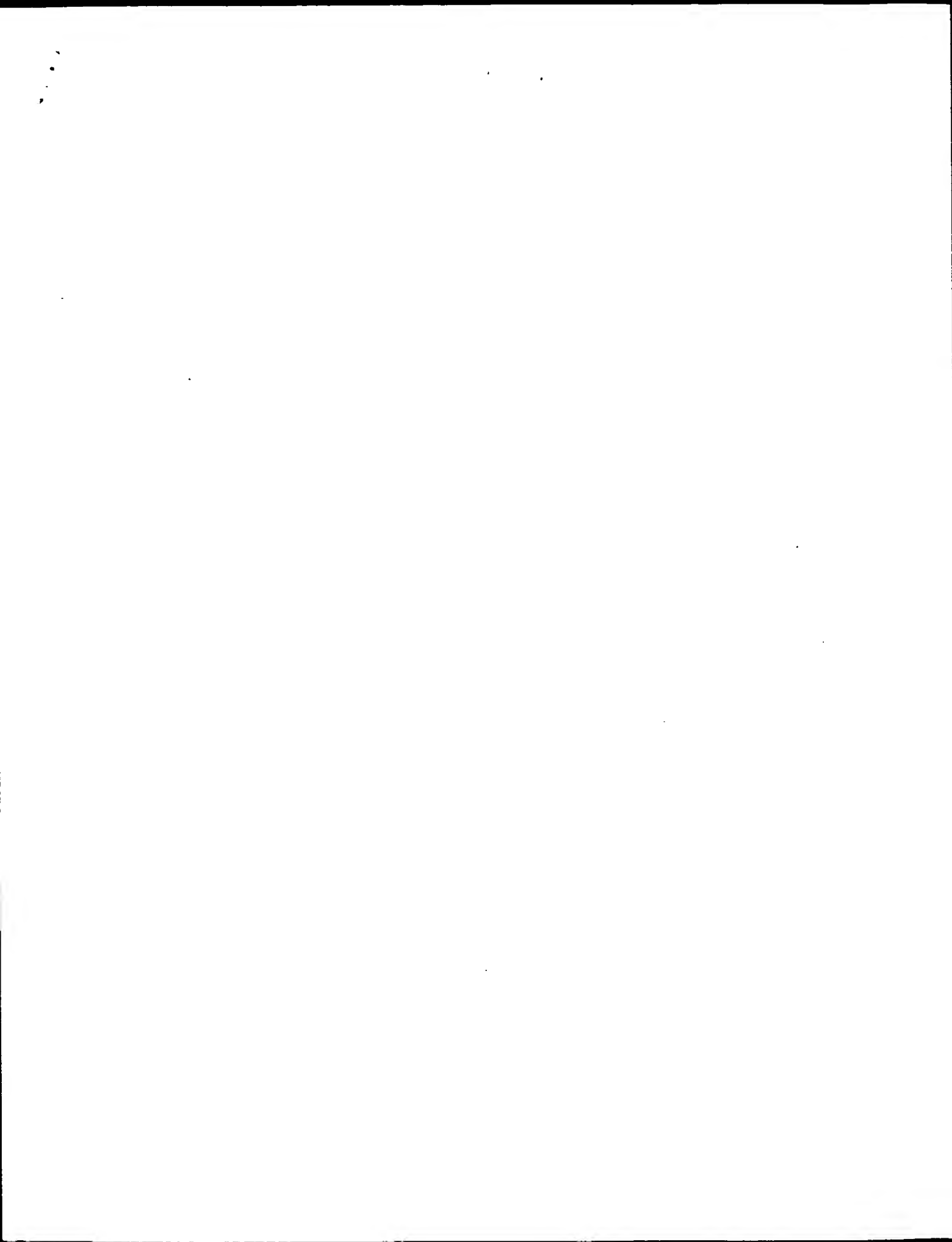
clone_end:77

vector_side:left"

71099..112105

/note="assembly_fragment:04906

BASE COUNT	46693	a	37639	c	39101	g	52617	t	200	others
ORIGIN										
Fragment Chain: 1										
msc_feature										
112206..176250										
/note="assembly: fragment:02016.0"										
Query Match	12.1%	Score	201.8;	DB	2;	Length	176250;			
Best Local Similarity	61.1%	Pred.	No. 2.6e-33;							
Matches	367;	Conservative	0;	Mismatches	222;	Indels	12;	Gaps	2;	
Db 476	ggatctctcttcagcgggagcccccagctctgtgagcagcagcagcagacatctcttgacc	535								
Db 114328	ggctccacacacacacccgaaacccggccacctcagagagacagcagcagctcttgcgac	114387								
Qy 536	aacgcgtctcttagaccacacatgacaccac---accgcgtgctgacacacagcctcaccaagg	592								
Db 114388	aggccctattaggacagcgtctccgccacacagccacccatccacagctcagctcagatcggg	114447								
Qy 593	ctgtgtgcgtatccagcccaaggtgtgtccacatctccagccgcttgcagcttaaggcccgcg	652								
Db 114448	tcacatccggacccacgcccaagcag-----ctgcctgctgaacacattgaacgctctct	114498								
Qy 653	tcaccaccagctgacacagcactctctgtgtgcgagcgctgtgggaaagtgaatgca	712								
Db 114499	tgaagagagcaccacccacccacacagcacaagttcattctcgaaacatctggcacaattgac	114558								
Qy 713	agggaggtgtgatactcccccggaacgtttgcctctctgtcgtggtctgcacacagagtgccgt	772								
Db 114559	gagaggttaccggcccccggccgctctccctctctgtctgctgacccgtgacacgtgctct	114618								
Qy 773	gcttagcgcacagctctgtgtcaactatgacagctgcatgtgtgtgtgtgtgcagagcattct	832								
Db 114619	gctccggagagcattgtgtgaattacgggacctgcattgtgctgacacgtgacaggtttct	114678								
Qy 833	accacttcagaaatagagacagatagagggtcctgtgctgacacacccctgtctgtctcc	892								
Db 114679	acgcactctccaaatgattgattgagaggttcttacttcgaaataccatgcctctgttcac	114738								
Qy 893	gtctcaactctgtgcgcccgcgtgtctctcaatgagtgactctccgtgtgtgtgtgtcctgc	952								
Db 114739	agtcctcactctgctctccaaatgactgtgtgattggagccctgtcttctgtgcttaccctgct	114798								
Qy 953	tgcctctctaccgtctgtccacacggcgtgtgaaagcttgcgcccacgctgtgatacagcgc	1012								
Db 114799	tgcctctctaccctcccgccaaaggctgtctgaactgtgtgagggcctgtttacacgttga	114858								
Qy 1013	tgcgcgcgcctgtgtgtgcgctgtcaaacacacgaaagcgtatctgtcaaacagcagcagc	1072								
Db 114859	cccaccccccctgcgctcagagtgagaaactccaaacactgctattgtgaactgcagacgct	114918								
Qy 1073	g 1073									
Db 114919	c 114919									
RESULT 10										
AF176904	1889	bp	mRNA	linear	vnt	23-DEC-1999				
LOCUS	AF176904									
DEFINITION	Gallus gallus sprouty 2 (spry2) mRNA, complete cds.									
ACCESSION	AF176904									
VERSION	AF176904.1									
KEYWORDS										
SOURCE	chicken.									
ORGANISM	Gallus gallus									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
AUTHORS	Archosauaria; Aves; Neognathae; Galliformes; Phasianidae;									
	Phasianinae; Gallus.									
	1 (bases 1 to 1889)									
	Minowada, G., Jarvis, L.A., Chi, C.L., Neubuser, A., Sun, X.,									
	Hachohen, N., Krasnow, M.A. and Martin, G.R.									
TITLE	Vertebrate sprouty genes are induced by FGF signaling and can cause									
	chondrodysplasia when overexpressed									



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:46:18 ; Search time 24.31 seconds
(without alignments)

1799.071 Million cell updates/sec

Title: US-10-082-902-2

Perfect score: 1657

Sequence: 1 MEPIPOSAPLTNSVWVOP.....SVICKASGDKTSRDKPF 299

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_17:

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhc:**
- 8: sp_organalle:**
- 9: sp_phage:**
- 10: sp_plant:**
- 11: sp_rodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1657	100.0	322	4 Q9C004	Q9C004 homo sapien
2	520	31.4	106	4 Q9C003	Q9C003 homo sapien
3	154.5	9.3	409	5 Q9V756	Q9V756 drosophila
4	126	7.6	738	11 Q88424	Q88424 mus musculus
5	125.5	7.6	984	5 Q9Y1P7	Q9Y1P7 cryptospori
6	121.5	7.3	757	5 Q9V2F2	Q9V2F2 drosophila
7	119.5	7.2	1365	4 Q96028	Q96028 homo sapien
8	116	7.0	1187	2 Q49549	Q49549 mycoplasma
9	112.5	6.8	470	4 Q9Y4Z3	Q9Y4Z3 homo sapien
10	112.5	6.8	617	4 Q9Y4Z4	Q9Y4Z4 homo sapien
11	111.5	6.7	723	11 Q9JKS4	Q9JKS4 mus musculus
12	111.5	6.7	723	11 Q9VWH2	Q9VWH2 mus musculus
13	111.5	6.7	734	4 Q75112	Q75112 homo sapien
14	111.5	6.7	832	4 Q9C010	Q9C010 homo sapien
15	111	6.7	867	5 Q9XTK3	Q9XTK3 giardia lam
16	109	6.6	432	4 Q9NPW2	Q9NPW2 homo sapien
17	109	6.6	494	4 Q95965	Q95965 homo sapien
18	109	6.6	898	4 Q9UF24	Q9UF24 homo sapien
19	109	6.6	4599	11 Q9J118	Q9J118 mus musculus

20	108.5	6.5	661	11 Q9JKS3	Q9JKS3 mus musculus
21	108	6.5	495	5 Q9C043	Q9C043 giardia lam
22	108	6.5	4599	4 Q9N2R2	Q9N2R2 homo sapien
23	107.5	6.5	4123	4 Q75851	Q75851 homo sapien
24	107	6.5	432	5 Q22884	Q22884 caenorhabdi
25	107	6.5	496	5 Q9VMD4	Q9VMD4 drosophila
26	107	6.5	4823	13 Q93321	Q93321 fugu rubrip
27	106.5	6.4	1119	5 Q18034	Q18034 caenorhabdi
28	106.5	6.4	1513	5 Q17970	Q17970 caenorhabdi
29	106.5	6.4	1616	5 Q9VSJ0	Q9VSJ0 drosophila
30	106	6.4	721	13 Q91902	Q91902 xenopus lae
31	106	6.4	5374	11 Q99ND0	Q99ND0 mus musculus
32	105.5	6.4	1704	5 Q94446	Q94446 chironomus
33	105.5	6.4	1713	11 Q88349	Q88349 mus musculus
34	105.5	6.4	3857	11 Q88840	Q88840 mus musculus
35	105	6.3	1534	4 Q75093	Q75093 homo sapien
36	105	6.3	2146	5 Q9VC97	Q9VC97 drosophila
37	104.5	6.3	1587	4 Q9Y6N6	Q9Y6N6 homo sapien
38	104	6.3	403	5 Q18375	Q18375 drosophila
39	104	6.3	564	4 Q9U1L7	Q9U1L7 homo sapien
40	104	6.3	1664	5 Q9TVQ2	Q9TVQ2 caenorhabdi
41	103.5	6.2	806	5 P91808	P91808 strongyloce
42	103	6.2	734	10 Q9LNZ0	Q9LNZ0 arabidopsis
43	103	6.2	1420	5 Q9GZJ3	Q9GZJ3 leishmania
44	103	6.2	1511	4 Q75412	Q75412 homo sapien
45	103	6.2	2872	11 Q9WUH8	Q9WUH8 rattus norv

ALIGNMENTS

```

RESULT 1
ID Q9C004 PRELIMINARY; PRT; 322 AA.
AC Q9C004;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE SPROUTY-4A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 111
RP SEQUENCE FROM N.A.
RC TISSUE=UMBILICAL ARTERY;
RA Leekema O.C., van Achterberg T.A.E., Spaargaren M.,
RA Kr von dem Borne A.E.G., Pannekoek H., de Vries C.J.M.;
RT *Identification of a novel human Sprouty homolog.*;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF227516; AAK00652.1;
SQ SEQUENCE 322 AA; 34929 MW; 5BF5B67854CDD6DD CRC64;

```

Query Match	100.0%	Score 1657;	DB 4;	Length 322;
Best Local Similarity	100.0%	Pred. No. 1.2e-142;		
Matches 299;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			
Qy 1	MEPIPOSAPLTNSVWVOP	LDLSRMSRSRLQHP	LTLPIDQVKTSHVENDY	IDNPSLAL 60
Db 24	MEPIPOSAPLTNSVWVOP	LDLSRMSRSRLQHP	LTLPIDQVKTSHVENDY	IDNPSLAL 83
Qy 61	TTGPKRTRGAP	ELAPTPARCDQDVTHWISFSGRPSVSSSSSTSSDQRLDHPAPPV	120	
Db 84	TTGPKRTRGAP	ELAPTPARCDQDVTHWISFSGRPSVSSSSSTSSDQRLDHPAPPV	143	
Qy 121	ADQASPRVRIQPKV	YVHCOPLDKGPVPPDLKHELCEACGCKCKEASPTLPSCW	180	
Db 144	ADQASPRVRIQPKV	YVHCOPLDKGPVPPDLKHELCEACGCKCKEASPTLPSCW	203	
Qy 181	VCNQECICSAQTLVNYGT	CMCLVQGIYFVHCTNEDDEGSCADHPCCSRNCCARWFSFMA	240	
Db 204	VCNQECICSAQTLVNYGT	CMCLVQGIYFVHCTNEDDEGSCADHPCCSRNCCARWFSFMA	263	

```

QY 241 LSVVPLCLCYLPATCGVKLAORGVDRLRRRCCKKHTNSVICKAASGDARTSRPKPF 299
DB 264 LSVVPLCLCYLPATCGVKLAORGVDRLRRRCCKKHTNSVICKAASGDARTSRPKPF 322

RESULT 2
ID Q9C003 PRELIMINARY; PRT; 106 AA.
AC Q9C003;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE SPROUTY-4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UMBILICAL ARTERY;
RA Leeksmä O.C., van Achterberg T.A.E., Spaargaren M.,
RA Kr von dem Borne A.E.G., Pannekoek H., de Vries C.J.M.;
RT "Identification of a novel human Sprouty homolog."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF227517; AAK00653.1; -
SQ SEQUENCE 106 AA; 11482 MW; 6158571E3477A598 CRC64;

Query Match 31.4%; Score 520; DB 4; Length 106;
Best Local Similarity 97.0%; Pred. No. 6.1e-40;
Matches 97; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEPPFQSAPLTPNSVWVQPLDLSRMSRLQHPILTLIPDQVKTSHVENDYIDNPSLAL 60
DB 1 MEPPFQSAPLTPNSVWVQPLDLSRMSRLQHPILTLIPDQVKTSHVENDYIDNPSLAL 60

QY 61 TTGPKTRGAGELAPATPARCDDVTHHWISFSGRPSVS 100
DB -61 TTGPKTRGAGELAPATPARCDDVTHHWISFSGRPSAT 100

RESULT 3
ID Q9V756 PRELIMINARY; PRT; 409 AA.
AC Q9V756;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CG10155 PROTEIN.
GN CG10155.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Cealniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.O., Helt G., Nelson C.R., Miklos G.L.G.,
RA Adair J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

```

```

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckl A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulo G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AF003813; AAF58207.1; -
DR FlyBase; FBgn0032964; CG10155.
DR InterPro; IPR000697; RanBPL_WASP.
DR InterPro; IPR001960; WH1.
DR SMART; SM00461; WH1; 1.
SQ SEQUENCE 409 AA; 44311 MW; 6FC0BD89FF3F9C96 CRC64;

Query Match 9.3%; Score 154.5; DB 5; Length 409;
Best Local Similarity 23.2%; Pred. No. 3.2e-06;
Matches 66; Conservative 22; Mismatches 99; Indels 97; Gaps 14;

QY 52 YIDNPSLALTGTGPKTRGAP--ELATPARCQ-----DVTHHWISF----- 92
DB 176 YISTDKTSATTSP-----DAPPASAPSPATAGIAASENYSVTLTAVHHIDYNPVVDQ 231

QY 93 -----SGRPSSVSSSTSSDQRLDHMAPPPVADQASPRVRIQPKVHCQ-----L 141
DB 232 PVGAQVNLNARRESIALKRNAL-----AAQAAAAQAAGGLACRDGGSGK 280

QY 142 DLKGPVPPPELDKHF-LICEAC-----GKCKCKECASPTLPCSWVCNQCCLCS 189
DB 281 PLHKPNVSDILKKTETRLCRYCHELYSEDFNRRGACE-----YAPDAFRSGYECIS- 331

QY 190 AQTLYNYGTCMLVQGIYFHTCTNEDDEGSCADHPSCSRS--NCCARMSFGALSVVLP 247
DB 332 -----GMCARCMI-----YHCMS-DABGTAHQPCDCAEAGCKRMLGLAILSLFVPC 381

QY 248 ILCYLPATCGVKLAORGVDRLRRRCCKKHTNSVICKAASGDARTSRPKPF 291
DB 382 LWCYPLRAC-----HLAGTHGCLGCGGQHK 406

RESULT 4
ID O88424 PRELIMINARY; PRT; 738 AA.
AC O88424;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE PACTOLUS.
GN ITGB2L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```


354 - NKHKQSVCHUNG-----PCYPCKLESINCRC--GTTKKSVP 398

RESULT 7

OS96028 PRELIMINARY; PRT; 1365 AA.

ID O96028

AC O96028;

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE WHSCL PROTEIN.

GN WHSCL OR MMSET OR TRX5.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI_TaxID=9606;

PN [1]

RP SEQUENCE FROM N.A.

RR MEDLINE=98282232; PubMed=9618163;

RRX Stec I., Wright T.J., Van Ommen G.J.B., de Boer P.A.J.,

RA van Haeringen A., Moorman A.F.M., Alther M.R., Den Dunnen J.T.;

RA "WHSC1, a 90 kb SET domain-containing gene, expressed in early

RR development and homologous to a Drosophila dysmorph gene maps in the

RRT wolf-hirschhorn syndrome critical region and is fused to IgH in

RRT t(4;14) multiple myeloma.";

RR Hum. Mol. Genet. 7:1071-1082(1998).

RL [2]

RP SEQUENCE FROM N.A.

RR TISSUE=MYELOMA;

RRX MEDLINE=99005326; PubMed=9787135;

RA Chesi M., Nardini E., Lim R.S., Smith K.D., Kuehl W.M.,

RA Bergsagel P.L.;

RA "The t(4;14) translocation in myeloma dysregulates both FGFR3 and a

RR novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts.";

RR Blood 92:3025-3034(1998).

RL [3]

RP SEQUENCE FROM N.A.

RR Angrand P.O., Valvatne H., Jeanmougin F., Adamson A.,

RA van der Hoeven F., Olsen L., Tekotte H., Huang N., Poch O.,

RA Lamerding J., Chambon P., Lessner R., Stewart A., Aasland R.;

RA "Mammalian Trichorax- and ASH1-like proteins: Putative chromatin

RR regulators which contain PHD fingers and SET domains.";

RR Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

RL [4]

RP SEQUENCE FROM N.A.

RR Chesi M., Bergsagel P.L.;

RRX Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

RRL EMBL; AF083386; AAD19343.1; -

RRR EMBL; AF071593; AAC24150.1; -

RRR EMBL; AJ007042; CAB45386.1; -

RRR EMBL; AF083387; AAD21770.1; -

RRR EMBL; AF083388; AAD21771.1; -

RRR EMBL; AF178219; AAF23370.1; -

RRR EMBL; AF178198; AAF23370.1; JOINED.

RRR EMBL; AF178199; AAF23370.1; JOINED.

RRR EMBL; AF178200; AAF23370.1; JOINED.

RRR EMBL; AF178201; AAF23370.1; JOINED.

RRR EMBL; AF178202; AAF23370.1; JOINED.

RRR EMBL; AF178203; AAF23370.1; JOINED.

RRR EMBL; AF178204; AAF23370.1; JOINED.

RRR EMBL; AF178205; AAF23370.1; JOINED.

RRR EMBL; AF178207; AAF23370.1; JOINED.

RRR EMBL; AF178208; AAF23370.1; JOINED.

RRR EMBL; AF178209; AAF23370.1; JOINED.

RRR EMBL; AF178210; AAF23370.1; JOINED.

RRR EMBL; AF178211; AAF23370.1; JOINED.

RRR EMBL; AF178212; AAF23370.1; JOINED.

RRR EMBL; AF178213; AAF23370.1; JOINED.

RRR EMBL; AF178214; AAF23370.1; JOINED.

RRR EMBL; AF178215; AAF23370.1; JOINED.

RRR EMBL; AF178216; AAF23370.1; JOINED.

RRR EMBL; AF178217; AAF23370.1; JOINED.

RRR EMBL; AF178218; AAF23370.1; JOINED.

DR InterPro: IPR000313; PMP.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000910; HMG_12_box.
 DR InterPro: IPR001214; SET.
 DR InterPro: IPR001841; Znf_ring.
 DR InterPro: IPR001965; PHD.
 DR InterPro: IPR003616; PostSET.
 DR Pfam: PF00628; PHD; 3.
 DR Pfam: PF00855; PMP; 2.
 DR Pfam: PF00856; SET; 1.
 DR PROSITE: PS50280; SET; 1.
 DR SMART: SM00398; HMG; 1.
 DR SMART: SM00398; HMG; 1.
 DR SMART: SM00249; PHD; 4.
 DR SMART: SM00508; PostSET; 1.
 DR SMART: SM00293; PMP; 2.
 DR SMART: SM00184; RING; 2.
 DR SMART: SM00317; SET; 1.
 SQ SEQUENCE 1365 AA; 152257 MW; 7B3128E1FA893AAA CRC64;

Query Match 7.2%; Score 119.5; DB 4; Length 1365;
 Best Local Similarity 23.7%; Pred. No. 0.015;
 Matches 55; Conservative 25; Mismatches 79; Indels 73; Gaps 12;

QY 97 SSVSSSTSSDRLDLMAPP-----VADQSPRAVRIQPVHCOF 140
 DB 614 SKSSPSALTEENVSDSPGDESESPYSADEQTQEVSVSSKSEKVTAKKEYV-CQL 672
 QY 141 LDLAGRAVPELDKHLCE-----AC-----GKCKCEKCSPTTDFSCWCM 183
 DB 673 CEKPG-----SLLCGPGCCGAFHLACGLSRREGRPTSECS--GHSCFVCK 721
 QY 184 QECLSAQTLVNTGTCMCLVQGLFYH--CTNE-----DDEG-SCADHPC-SCSRNCCA 233
 DB 722 -----ESKTDVRCVVCVQCGRFHEACVKKPLTVFESRGFRPLHSCSHASPSN 774
 QY 224 RMSPGALSVLPCLLCYLPATGCVKLAORGDRLRRCCKKHNSTYCA 285
 DB 775 PRPSKGMKRCVRCPAVHNSGACL-----AAGCVYASNSITCTA 815

RESULT 8
 Q49549 PRELIMINARY; PRT: 1187 AA.
 AC Q49549;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE REPEAT REGIONS IN POTENTIAL METAL BINDING PROTEIN GENE REGION.
 OS Mycoplasma hyorhinis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID-2100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GD1;
 RX MEDLINE-95014025; PubMed-7928953;
 RA Deng G., McIntosh M.A.;
 RT "An amplifiable DNA region from the Mycoplasma hyorhinis genome.";
 RL J. Bacteriol. 176:5929-5937(1994).
 DR EMBL: L11447; AAA62228.1;
 SQ SEQUENCE 1187 AA; 135027 MW; 66A82A5B0ED93E CRC64;

Query Match 7.0%; Score 116; DB 2; Length 1187;
 Best Local Similarity 21.9%; Pred. No. 0.027;
 Matches 59; Conservative 25; Mismatches 113; Indels 72; Gaps 14;
 QY 35 LTILPIDOVKTSVENDYIDNPISLALTGPKRTRGGAPAPLAPTRACD--ODVTHHMTSF 92
 DB 266 LTDLDDHFEIVHLEEV-----CLA-----COHVATCDICKNLSSELY 306

QY 93 SGRPSVSSSSSTSSDRLDLMAPPVADQSPRAVRIQPVHCOPLDKGPAVPEL 152
 DB 307 RLKNGQVNLLETELETEL-HYADSPQEGKEPGCSLKRTESCSO----- 352
 QY 153 DKHFLCEAGCKKCKEC-----ASPTL-----PSCWNCDELCSQTLV-NYTCMKL 202
 DB 353 -----CEA-----CKCECEKSCSELTCGCEATCSADHCCGCESCACPNTYCACT 403
 QY 203 VQGLFYHCTNEDDEGSCADHPCSCSRNCCARWSPGALSVLPCLLCYLPV-----TCGV 258
 DB 404 EE-----HCECTESTGCENEPCEPEBACDSEHCECEVDETQACLDGNTQADTKVCGGT 459
 QY 259 KLAORGDRLRRCG-CKKHTNSV-ICKA 285
 DB 460 -----OEHPTECEKECECKCKCKA 480

RESULT 9
 Q9Y423 PRELIMINARY; PRT: 470 AA.
 AC Q9Y423;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ZASP PROTEIN (FRAGMENT).
 GN ZASP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SKELETAL MUSCLE, AND HEART;
 RA Faulkner G., Pallavicini A., Formentin E., Comelli A., Ievolella C.,
 RA Trevisan S., Bortolotto G., Scannapieco P., Salamon M., Mouly G.,
 RA Valle G., Lanfranchi G.;
 RT "ZASP, a new z-band alternatively spliced PDZ-motif protein.";
 RL J. Cell Biol. 0:0-0(0).
 CC 1- SIMILARITY: CONTAINS LIM DOMAIN(S). THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 CC EMBL: AJ133768; CAB46729.1;
 DR HSSP: Q05158; IOLI.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR001781; LIM.
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR002965; P-rich_extensn.
 DR Pfam: PF00412; LIM; 3.
 DR Pfam: PF00595; PDZ; 1.
 DR PRINTS: PRO1217; PRICHEXTENS.
 DR PRODOM: PD000094; LIM; 3.
 DR SMART: SM00132; LIM; 3.
 DR SMART: SM00228; PDZ; 1.
 DR PROSITE: PS00280; IG_MHC; UNKNOWN.1.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE: PS0023; LIM_DOMAIN_2; 3.
 DR PROSITE: PS50106; PDZ; 1.
 DR LIM motif; Metal-binding; zinc.
 FT NON_TER 1
 SQ SEQUENCE 470 AA; 50661 MW; AE787648FA1A2BCF CRC64;

Query Match 6.8%; Score 112.5; DB 4; Length 470;
 Best Local Similarity 21.3%; Pred. No. 0.023;
 Matches 77; Conservative 42; Mismatches 132; Indels 111; Gaps 19;
 QY 4 PIPGSAFLTPNSV---MVQLDSRMSHSRLQH---PLTILPIDOVKTSVENDYIDNP 56
 DB 134 PAPRPVTVTSIRPVSQVPASVSPGANYSPPTYPXPA-----YTSP 184
 QY 57 SLATTPKTRTRGGAPAPLAPTRACDODVTHHMTSFSGRPSVSS----- 101
 DB 185 APATTPSPVPTYPSPAPATTPSPAPNTPAPSVATSGPAPAPSPAPVTTDTSFSQKFA 244

QY 102 ---SSSTSSDQRLDHP-----PVADQASPAVRI-----QPKVHCOPDLJK 145
 DB 245 PGRSTTSSKQTLPRGPAATYPAGPOVPLARCTVGRARERPASSFTPLCGHCNNV-IRG 303
 QY 146 PAV-----PELDKHFLLCEAGCK-----CKEASPRLLPSCW 181
 DB 304 PELVAMGRSMHDEEP-----TCAYCKTSLADVCVEEONNVYCERCYEQGFAPLCAK 355
 QY 182 CNOECLCSAOTLVN---YGTG-MCLV-----OGIFYHCTNEDDEGSCADHPCSCRSNC 231
 DB 356 CNTKINGEVMALROTHWTTCFVCAACKKPFNSLFH---MEDGEYCEKDYINLFSTKC 412
 QY 232 --C-----ARMSFGAL-----SVLPCLLCYLPATGCVKLAORGYDLRRPGCRKHTN 279
 DB 413 HGGDFEVEAGDKFIEALGHTWHDTCFCAVCHVNLG-----QPFYSKDRPLCK-KHAN 466
 QY 280 SV 281
 DB 467 TI 468

RESULT 10
 QY424 PRELIMINARY; PRT; 617 AA.

AC QY424;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ZASP PROTEIN (FRAGMENT).
 GN ZASP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SKELETAL MUSCLE AND HEART;
 RA Faulkner G., Fallavicioli A., Formentin E., Comelli A., Ievoliella C.,
 RA Trevisan S., Bortoletto G., Scannapieco P., Salamon M., Mouly G.,
 RA Valle G., Lantfranchi G.;
 RT "ZASP: a new 2-band alternatively spliced PDZ-motif protein.";
 RL J. Cell Biol. 0:0-0(0).
 CC -1- SIMILARITY: CONTAINS LIM DOMAIN(S). THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 DR EMBL: AJ133767; CAB6728.1; -
 DR InterPro: IPR000345; CytoC_heme_bind.
 DR InterPro: IPR001781; LIM.
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR002965; P-rich_extensn.
 DR Pfam: PF00412; LIM; 3.
 DR Pfam: PF00595; PDZ; 1.
 DR PRINTS: PRO1217; PRICHEXTENSN.
 DR ProDom: PD000094; LIM; 3.
 DR SMART: SM00132; LIM; 3.
 DR SMART: SM00228; PDZ; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS00478; LIM_DOMAIN_1; UNKNOWN_1.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 3.
 DR PROSITE: PS50106; PDZ; 1.
 DR LIM motif; Metal-binding; zinc.
 KW NON_TER 1
 FT SEQUENCE 617 AA: 66518 MW: 22FC77402352A340 CRC64;

Query Match 6.8%; Score 112.5; DB 4; Length 617;
 Best Local Similarity 21.3%; Pred. No. 0.03;
 Matches 77; Conservative 42; Mismatches 132; Indels 111; Gaps 19;

QY 4 PLFQSAFLTPNSV---WVQPLDSRMSHSLQH----PLTILPIDQVTSHEVDYNDP 56
 DB 281 PPKPRRVVTSIRPSVYQVPAVSTYSPSGANYSPTPTSPAPAA-----YTPSP 331
 QY 57 SLATLTGKRTGCAPELAPFARCDQDVTHHWISFGRRPSSVSS----- 101

DB 332 APAYTPSPVPTTSPAPAVTTPAPNPAPSVASGAPAPASRPWYTDTSFQKFA 391
 QY 102 ---SSSTSSDQRLDHP-----PVADQASPAVRI-----QPKVHCOPDLJK 145
 DB 392 PGRSTTSSKQTLPRGPAATYPAGPOVPLARCTVGRARERPASSFTPLCGHCNNV-IRG 450
 QY 146 PAV-----PELDKHFLLCEAGCK-----CKEASPRLLPSCW 181
 DB 451 PELVAMGRSMHDEEP-----TCAYCKTSLADVCVEEONNVYCERCYEQGFAPLCAK 502
 QY 182 CNOECLCSAOTLVN---YGTG-MCLV-----OGIFYHCTNEDDEGSCADHPCSCRSNC 231
 DB 503 CNTKINGEVMALROTHWTTCFVCAACKKPFNSLFH---MEDGEYCEKDYINLFSTKC 559
 QY 232 --C-----ARMSFGAL-----SVLPCLLCYLPATGCVKLAORGYDLRRPGCRKHTN 279
 DB 560 HGGDFEVEAGDKFIEALGHTWHDTCFCAVCHVNLG-----QPFYSKDRPLCK-KHAN 613
 QY 280 SV 281
 DB 614 TI 615

RESULT 11
 QY424 PRELIMINARY; PRT; 723 AA.

AC QY424;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ORACLE 1 PROTEIN (FRAGMENT).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-HEART;
 RA MEDLINE-20193503; PubMed-10727866;
 RA Passier R., Richardson J.A., Olson E.N.;
 RT "Oracle, a novel PDZ-LIM domain protein expressed in heart and
 RT skeletal muscle.";
 RL Mech. Dev. 92:277-284(2000).
 CC -1- SIMILARITY: CONTAINS LIM DOMAIN(S). THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 DR EMBL: AF228057; AAF3847.1; -
 DR InterPro: IPR000345; CytoC_heme_bind.
 DR InterPro: IPR003006; IQ_MHC.
 DR InterPro: IPR001781; LIM.
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR002965; P-rich_extensn.
 DR Pfam: PF00412; LIM; 3.
 DR Pfam: PF00595; PDZ; 1.
 DR PRINTS: PRO1217; PRICHEXTENSN.
 DR ProDom: PD000094; LIM; 3.
 DR SMART: SM00132; LIM; 3.
 DR SMART: SM00228; PDZ; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS00290; IQ_MHC; UNKNOWN_1.
 DR PROSITE: PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE: PS50023; LIM_DOMAIN_2; 3.
 DR PROSITE: PS50106; PDZ; 1.
 DR LIM motif; Metal-binding; zinc.
 KW NON_TER 723
 FT SEQUENCE 723 AA: 76431 MW: CC67D38AC2FFA6B6 CRC64;

Query Match 6.7%; Score 111.5; DB 11; Length 723;
 Best Local Similarity 19.5%; Pred. No. 0.043;
 Matches 78; Conservative 47; Mismatches 125; Indels 149; Gaps 22;

QY 1 MEPIPSAFLTPNSVWVQPLDSRMSHSLQHPLTILPIDQVNT----- 45

```

Db 354 VENNROQASVSP-AAAASPAASHTSY-S-EGPAAPAKRVYTTASIRPSYQOPVPAS 410
Qy 46 -----SHVENDYIDNPSLALT-----TGPKRTGAGAPLAPTPA 79
Db 411 SYSPSPGANSPTPTPPSPAPAYTPPAPAYTPPAPAYTPPAPAYTPPAPAYTPPS 470
Qy 80 RCDODVTHMVISFGRSPSSVS-----SSSTSSDORLLDHMAPP--PV 120
Db 471 -----AAVSGSPSESASRPWWYDSDFSQKFAFGKSTTTVSKOTLPAGAPAYPT 520
Qy 121 ADQASP-----RAVRI-----QPKVHCQPLDKGPAV-----PELDKHFLLCE 160
Db 521 GPQVYPLANGTFORARERFPASSRTPLCGHCHNNV-IRGFLVAMGRSMHPEEFN----- 572
Qy 161 ACCKCK-----CKEASPRTLPSQWVNOECLSAQTLVN---YGTC-M 200
Db 573 -CAVCKTSLADVCYFEQNNVYCEQYQDFAPICACNKRKINGEVMHMLRQTHHTTCV 631
Qy 201 CLV-----QGIFYHCTNDDDEGSCADHPCGCSNSNC--C-----ARMSFGAL-----S 242
Db 632 CAACKKPFGENSLFH--MEDGEPYCEKDYINLFSTKCHGCDPEVEAGDKFIEALGHTWHD 688
Qy 243 VWPLCLCYLPATGCVKLAQNGYDRLRRPGCRCKHTNSV 281
Db 689 TCFICAVCHVNLEG-----QPFYSKKDKPLCK-KHMAI 721

```

RESULT 12

```

ID 09VWH2 PRELIMINARY; PRT: 723 AA.
AC 09VWH2;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CYPHER1.
GN zasp.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=99321910; PubMed=10391924;
RA Zhou Q., Ruiz-Lozano P., Martone M.E., Chen J.;
RT Cypher, a Striated Muscle-restricted pdz and LIM Domain-containing
RT protein, binds to alpha-Actinin-2 and Protein Kinase C.;
RL J. Biol. Chem. 274:19807-19813(1999).
RN 12
RP SEQUENCE FROM N.A.
RA Chen J., Zhou Q.;
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS LIM DOMAIN(S). THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR EMBL: AF114378; AADA2950.2; -
DR HSSP: 005158; IOLI.
DR MGD: MGI:134412; zasp.
DR InterPro: IPR000345; CYC_heme_bld.
DR InterPro: IPR001781; LIM.
DR InterPro: IPR001478; PDZ.
DR InterPro: IPR002965; P_rich_extensn.
DR Pfam: PF00412; LIM; 3.
DR PRINTS: PR01217; PRICHEXTENS.
DR PRODOM: PD000094; LIM; 3.
DR SMART: SM00132; LIM; 3.
DR SMART: SM00228; PDZ; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS50023; LIM_DOMAIN_2; 3.
DR PROSITE: PS50106; PDZ; 1.
KW LIM motif; Metal-binding; Zinc.
SQ SEQUENCE 723 AA; 76445 MW; CC66829FD2PD0691 CRC64;

```

Query Match 6.7%, Score 111.5; DB 11; Length 723;
 Best Local Similarity 19.5%, Pred. No. 0.043;
 Matches 78; Conservative 47; Mismatches 125; Indels 149; Gaps 22;

```

Qy 1 MEPRQASAPLVNWMVQPLDSSMSHRLDHLTLPLDOYKT----- 45
Db 354 VENNROQASVSP-AAAASPAASHTSY-S-EGPAAPAKRVYTTASIRPSYQOPVPAS 410
Qy 46 -----SHVENDYIDNPSLALT-----TGPKRTGAGAPLAPTPA 79
Db 411 SYSPSPGANSPTPTPPSPAPAYTPPAPAYTPPAPAYTPPAPAYTPPAPAYTPPS 470
Qy 80 RCDODVTHMVISFGRSPSSVS-----SSSTSSDORLLDHMAPP--PV 120
Db 471 -----AAVSGSPSESASRPWWYDSDFSQKFAFGKSTTTVSKOTLPAGAPAYPT 520
Qy 121 ADQASP-----RAVRI-----QPKVHCQPLDKGPAV-----PELDKHFLLCE 160
Db 521 GPQVYPLANGTFORARERFPASSRTPLCGHCHNNV-IRGFLVAMGRSMHPEEFN----- 572
Qy 161 ACCKCK-----CKEASPRTLPSQWVNOECLSAQTLVN---YGTC-M 200
Db 573 -CAVCKTSLADVCYFEQNNVYCEQYQDFAPICACNKRKINGEVMHMLRQTHHTTCV 631
Qy 201 CLV-----QGIFYHCTNDDDEGSCADHPCGCSNSNC--C-----ARMSFGAL-----S 242
Db 632 CAACKKPFGENSLFH--MEDGEPYCEKDYINLFSTKCHGCDPEVEAGDKFIEALGHTWHD 688
Qy 243 VWPLCLCYLPATGCVKLAQNGYDRLRRPGCRCKHTNSV 281
Db 689 TCFICAVCHVNLEG-----QPFYSKKDKPLCK-KHMAI 721

```

RESULT 13

```

ID 075112 PRELIMINARY; PRT: 734 AA.
AC 075112;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE KIA0613 PROTEIN (FRAGMENT).
GN KIA0613.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.;
RL DNA Res. 5:169-176(1998).
CC -1- SIMILARITY: CONTAINS LIM DOMAIN(S). THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR EMBL: AB014513; BAA31588.1; -
DR HSSP: 005158; IOLI.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR001781; LIM.
DR InterPro: IPR001478; PDZ.
DR InterPro: IPR002965; P_rich_extensn.
DR Pfam: PF00412; LIM; 3.
DR PRINTS: PR01217; PRICHEXTENS.
DR PRODOM: PD000094; LIM; 3.
DR SMART: SM00132; LIM; 3.
DR SMART: SM00228; PDZ; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS00478; LIM_DOMAIN_1; 2.

```


• Sat May 11 12:51:27 2002

us-10-082-902-2.rspt

Page 9

Search completed: May 8, 2002, 10:49:36
Job time: 198 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:46:43 ; Search time 9.88 seconds
(without alignments)
1109.593 Million cell updates/sec

Title: US-10-082-902-2
Perfect score: 1657
Sequence: 1 MEPPIPQSAFLTPNSVAVNP.....SVYCAASGAKTSRPPKPF 299

Scoring table:

Searched: 100059 seqs, 3664827 residues
Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1657	100.0	299	SPY4_HUMAN	09C004 homo sapien
2	1554.5	93.8	300	SPY4_MOUSE	09WEP2 mus musculu
3	673.5	40.6	313	SPY2_CHICK	09P112 gallus gall
4	667	40.3	313	SPY1_MOUSE	09GXV9 mus musculu
5	641.5	38.7	315	SPY2_HUMAN	043597 homo sapien
6	633.5	38.2	315	SPY2_MOUSE	09GXV8 mus musculu
7	622.5	37.6	288	SPY3_HUMAN	043610 homo sapien
8	538	31.9	139	SPY1_HUMAN	043609 homo sapien
9	320	23.6	589	SPY1_MOUSE	044783 drosophila
10	250	15.1	76	SPY1_CHICK	09P111 gallus gall
11	130	7.2	2476	ZAN_PIG	044783 drosophila
12	113	6.8	809	PAT3_CAEEL	027874 caenorhabdi
13	110.5	6.7	2813	VWF_HUMAN	024275 homo sapien
14	108	6.5	2139	CRB_DROME	P10040 drosophila
15	107.5	6.5	1700	BAR3_CHITE	003376 chironomus
16	107	6.5	1712	TGFB_RAT	000818 rattus norv
17	106.5	6.4	2482	VWF_PIG	028633 sus scrofa
18	105	6.3	5376	ZAN_MOUSE	088799 mus musculu
19	104	6.3	363	PSPB_CANFA	P17129 canis faml
20	104	6.2	618	DLI3_HUMAN	09NY17 homo sapien
21	103.5	6.2	1972	P531_HUMAN	012888 homo sapien
22	102.5	6.2	577	ITB6_CAYPO	P18563 canis porce
23	102.5	6.2	4543	LBP1_CHICK	P98157 gallus gall
24	102	6.2	1799	LMB2_MOUSE	061292 mus musculu
25	101	6.1	2871	FBN1_MOUSE	061554 mus musculu
26	101	6.1	3149	TEGU_EBV	P03186 epstein-bar
27	100.5	6.1	771	ITB2_MOUSE	P11835 mus musculu
28	100	6.0	1964	NTC4_MOUSE	P31695 mus musculu
29	99	6.0	619	MTN4_HUMAN	095460 homo sapien
30	98	5.9	2444	NTC1_HUMAN	P46531 homo sapien
31	97.5	5.9	614	NRD1_HUMAN	P20127 ononis yell
32	97.5	5.9	798	ITB7_HUMAN	088671 rattus norv
33	97	5.9	1113	COR1_MOUSE	P32592 bos taurus

34	97	5.9	2700	ZAN_HUMAN	09Y493 homo sapien
35	96.5	5.8	2569	LMA3_MOUSE	061789 mus musculu
36	96.5	5.8	2871	FBN1_HUMAN	P35555 homo sapien
37	96.5	5.8	3110	LMA2_HUMAN	P24043 homo sapien
38	96	5.8	457	ODR7_CAEEL	P41933 caenorhabdi
39	96	5.8	965	YNC3_YEAST	P53971 saccharomyc
40	96	5.8	2813	VWF_CANFA	028295 canis faml
41	95.5	5.8	1776	POLR_OYAV	P20127 ononis yell
42	95.5	5.8	2531	NTC1_RAT	007008 rattus norv
43	95	5.7	589	ITB2_RAT	088671 rattus norv
44	95	5.7	769	ITB2_BOVIN	P32592 bos taurus
45	95	5.7	769	ITB2_HUMAN	P05107 homo sapien

ALIGNMENTS

RESULT 1
SPY4_HUMAN STANDARD: PRT: 299 AA.
ID 09C004; 09C003;
AC 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DR SPROUTY HOMOLOG 4 (SPRY-4).
GN SPY4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Ombilical artery;
RA Leekema O.C., van Achterberg T.A.E., Spaargaren M.,
von dem Borne A.E.G., Pannekoek H., de Vries C.J.M.;
RT Identification of a novel human Sprouty homolog *;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
ORGANOGENESIS (BY SIMILARITY).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND C; MAY BE
PRODUCED BY ALTERNATIVE SPLICING.
CC -!- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
THE PROTEIN TO THE MEMBRANE RUPTILES.
CC -!- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
or send an email to license@sdb.ch).
CC
CC EMBL: AF227516; AAK00652.1; ALT_INIT.
CC EMBL: AF227517; AAK00653.1; -
CC Developmental protein; Membrane; Alternative splicing.
CC KW DOMAIN 97 107
CC FT DOMAIN 159 283
CC FT VARSPLIC 97 106
CC FT VARSPLIC 107 299
CC FT VARSPLIC MISSING (IN ISOFORM C).
CC SQ SEQUENCE 299 AA; 32541 MW; 105f6f1b89f7b6c3 CRC64;

Query Match 100.0%; Score 1657; DB 1; Length 299;
Best Local Similarity 100.0%; Pred. No. 5.9e-118;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPPIPQSAFLTPNSVAVNP...LDLSRMSHSLQHPRLTPIDQVKSXVNDYINPISAL 60
DB 1 MEPPIPQSAFLTPNSVAVNP...LDLSRMSHSLQHPRLTPIDQVKSXVNDYINPISAL 60
QY 61 TTGPKRTGCAELAPTPARCDDQDVTHHWISFGSRSSVSSSSSTSSDORLLDHMAPPV 120

```

DB 61 TTGPKRTGGAPELAPTPARCDDVTHHMFSGRPSVSSSSSTSSDRLDHNAPPV 120
QY 121 ADOASPRARVIOPKVHVHCOPLDLKGAVPPELDKHFLLCEACGCKCKEASPRTPSPSC 180
DB 121 ADOASPRARVIOPKVHVHCOPLDLKGAVPPELDKHFLLCEACGCKCKEASPRTPSPSC 180
QY 181 VWCNOECICSAGTIVNVCMTLVGIFHYHCTNEDDESCADHPGCSRSNCCARMSFPGA 240
DB 181 VWCNOECICSAGTIVNVCMTLVGIFHYHCTNEDDESCADHPGCSRSNCCARMSFPGA 240
QY 241 LSVLPCLLCYLPATGCVKLAORGVDRLRRPGCRCKHTNSVLCKAASGDARTSRDPKPF 299
DB 241 LSVLPCLLCYLPATGCVKLAORGVDRLRRPGCRCKHTNSVLCKAASGDARTSRDPKPF 299

RESULT 2
SPY4_MOUSE STANDARD: PRT; 300 AA.
ID SPY4_MOUSE
AC 09WTF2; 09GXV7;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SPROUTY HOMOLOG 4 (SPRY-4).
GN SPRY4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX MEDLINE=99264295; PubMed=10330503;
RA de Maxly A.A., Nakatake Y., Moncada S., Itoh N., Thierry J.P.,
RA Bellusci S.;
RT *Cloning and expression pattern of a mouse homologue of drosophila
RT sprouty in the mouse embryo.*;
RL Mechn. Dev. 81:213-216(1999).
RN 121
RP SEQUENCE FROM N.A.
RX MEDLINE=99429807; PubMed=10498682;
RA Minowada G., Jarvis L.A., Chl C.L., Neubuser A., Sun X., Hacohen N.,
RA Krasnow M.A., Martin G.R.;
RT *Vertebrate sprouty genes are induced by FGF signaling and can cause
RT chondrodysplasia when overexpressed.*;
RL Development 126:4465-4473(1999).
CC -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
CC FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
CC ORGANOGENESIS.
CC -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS
CC BUT IS TRANSLOCATED TO THE MEMBRANE RUFFLES IN CELLS STIMULATED
CC WITH EGF (EPIDERMAL GROWTH FACTOR).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYO AND ADULT TISSUES
CC INCLUDING HEART, BRAIN, LUNG, KIDNEY, AND SKELETAL MUSCLE.
CC -1- DEVELOPMENTAL STAGE: AT 8 DPC EXPRESSED IN THE LATERAL PLATE
CC MESODERM OF THE PRIMITIVE STREAK. AT 9.5 AND 10.5 DPC EXPRESSED IN
CC THE NASAL PLACODES, MAXILLARY AND MANDIBULAR PROCESSES, POSTERIOR
CC PART OF THE HYOID ARCH AND THE PROGRESS ZONE OF THE LIMB BUDS AND
CC THE PRESOMITIC MESODERM. AT 11.5 DPC EXPRESSED IN THE DORSO-
CC LATERAL REGION OF THE SOMITES (MOSTLY IN THE MYOTOME) AND IN THE
CC OTIC VESICLE. AT 11.5 AND 12.5 DPC EXPRESSED IN THE DISTRAL LUNG
CC MESENCHYME, WITH A STRONG EXPRESSION IN THE ACCESSORY LOBE OF THE
CC LUNG.
CC -1- INDUCTION: BY FGF SIGNALING.
CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
CC THE PROTEIN TO THE MEMBRANE RUFFLES.
CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; AB019280; BAA7689.1; -
DR EMBL; AF176906; AAD56007.1; -
DR MGD; MGI:1345144; Spry4.
KW Developmental protein; Membrane.
FT DOMAIN 98 108 POLY-SER.
FT DOMAIN 160 284 CYS-RICH.
FT CONFLICT 10 10 V -> F (TN REF. 2).
SQ SEQUENCE 300 AA; 32523 MW; DA963036FCE0E73F CMC64;

Query Match 93.8%; Score 1554.5; DB 1; Length 300;
Best Local Similarity 92.7%; Pred. No. 3e-110;
Matches 278; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

QY 1 MEPPYPOSA-PLTPNSVWNPOLDSSRHSRIQHPITLPIQVYKTHVENDYIDNSLA 59
DB 1 MEPPYPOSSVPVNPSSVWVQPLDSRAPHSRIQHPITLPIQVYKTHVENDYIDNSLA 60
QY 60 LTTGPKRTGGAPELAPTPARCDDVTHHMFSGRPSVSSSSSTSSDRLDHNAPP 119
DB 61 PATGPKRRPGPELAPTPARCDDVTHHMFSGRPSVSSSSSTSSDRLDHNAPP 120
QY 120 ADOASPRARVIOPKVHVHCOPLDLKGAVPPELDKHFLLCEACGCKCKEASPRTPSPSC 179
DB 121 VAEQSPRAVRLQPKVHCKRPLDKGTPAPPELDKHFLLCEACGCKCKEASPRTPSPSC 180
QY 180 VWCNOECICSAGTIVNVCMTLVGIFHYHCTNEDDESCADHPGCSRSNCCARMSFPG 239
DB 181 VWCNOECICSAGTIVNVCMTLVGIFHYHCTNEDDESCADHPGCSRSNCCARMSFPG 240
QY 240 ALSVLPCLLCYLPATGCVKLAORGVDRLRRPGCRCKHTNSVLCKAASGDARTSRDPKPF 299
DB 241 ALSVLPCLLCYLPATGCVKLAORGVDRLRRPGCRCKHTNSVLCKAASGDARTSRDPKPF 300

RESULT 3
SPY2_CHICK STANDARD: PRT; 313 AA.
ID SPY2_CHICK
AC 09PTL2;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SPROUTY HOMOLOG 2 (SPRY-2).
GN SPRY2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
ON NCBI_TaxID=9031;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE=99429807; PubMed=10498682;
RA Minowada G., Jarvis L.A., Chl C.L., Neubuser A., Sun X., Hacohen N.,
RA Krasnow M.A., Martin G.R.;
RT *Vertebrate sprouty genes are induced by FGF signaling and can cause
RT chondrodysplasia when overexpressed.*;
RL Development 126:4465-4473(1999).
RN 121
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=20130796; PubMed=10662503;
RA Chambers D., Medhurst A.D., Walsh F.S., Price J., Mason I.;
RT *Differential display of genes expressed at the midbrain-hindbrain
RT junction identifies sprouty2: an fgfr8-inducible member of a family of
RT intracellular FGF antagonists.*;
RL Mol. Cell. Neurosci. 15:22-35(2000).
CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN FGF-MEDIATED PATTERNING OF
CC THE MID/HINDBRN REGION BY ACTING TO MODULATE THE SIGNALING
CC EFFECTS OF FGFR8 THROUGH PARTICIPATION IN A REGULATORY NEGATIVE
CC FEEDBACK LOOP.
CC -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS

```

CC BUT IS TRANSLOCATED TO THE MEMBRANE RUFFLES IN CELLS STIMULATED
 CC WITH EGF (EPIDERMAL GROWTH FACTOR).
 CC -1- TISSUE SPECIFICITY: BRAIN AND INTERLIMB REGION.
 CC -1- DEVELOPMENTAL STAGE: AT THE 4-TO 5-SOMITE STAGE (4/5S) FOUND IN
 CC THE EMBRYO IN SCATTERED CELLS ACROSS THE NEURAL PLATE IN THE
 CC PRESUMPTIVE MID/ HINDRAIN REGION. AT 7/8S FOUND IN THE ISTHUS
 CC AND THROUGHOUT THE PRESUMPTIVE RL TERRITORY. BETWEEN 10-14S STAGE
 CC FOUND THROUGHOUT THE RL REGION AND AT THE ISTHMIC CONstriction. BY
 CC 26S THE ANTERIOR LIMIT OF EXPRESSION EXTENDS INTO THE POSTERIOR
 CC MIDRAIN REGION AND THIS PATTERN OF EXPRESSION IS MAINTAINED AT
 CC LATER STAGES.
 CC -1- INDUCTION: BY FGF SIGNALING.
 CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
 CC THE PROTEIN TO THE MEMBRANE RUFFLES.
 CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF176904; AAD56005.1; -
 CC Developmental protein: Membrane.
 CC KW DOMAIN 176 299 CYS-RICH
 CC SEQUENCE 313 AA; 34417 MW; 614C335EDDC36EB3 CRC64;

Query Match 40.6%; Score 673.5; DB 1; Length 313;
 Best Local Similarity 44.8%; Pred. No. 5.7e-44;
 Matches 138; Conservative 39; Mismatches 76; Indels 55; Gaps 9;

QY 19 QPLDSMSHSR-----LQHPLTLPIDQVKTSHVENDYIDNPSLALTTGPKRTGR 68
 DB 12 QALQARDSRGRPELDLADVLQVHLSLDQIRALRNNEETEGTVAPRGVAK--- 68
 QY 69 GGAEPLATP-----ARCDDVTH-----HWISFGSPSSVSSSS 103
 DB 69 -SAPRSASQPKSERPHGLPEHRHFRGVQHTQTHSAPRAVLSISTVSTGSRSTRTSTS 127
 QY 104 STSDORLIDHMAPPVADQASPAVRIQPKVYHCOPLDKGPAPV---ELDKHFLLC 159
 DB 128 SNSSEQRLLG-SSSGPVAD---GIYRMPK-----SELSSSELPKSLKEDLAHSYRC 176
 QY 160 EACGKCKCKECASPTLLPSCWGNQECLSAQTLVNYGTGCMQVGIYFHCNTNDEGSC 219
 DB 177 EDCKKCKCKEETYPRLTPSCWICDKQCLCSAQNVVDYGTGVCCVYGLFYHCSN-DDEDNC 235
 QY 220 ADHPCGSRNSGCCARMSFGALSVVLPCLLYLPATGCVKLAQRGYDLRRPGCRCKHTN 279
 DB 236 ADNPGCSQSHCHCTKMSAMGVSLFLPLCMCTLPAGCKLKCQCGCTDVRNRPGRCKHNS 295
 QY 280 SVICKAAS 287
 DB 296 TVCKKVPSS 303

RESULT 4
 SPYL_MOUSE
 ID SPYL_MOUSE STANDARD; PRT; 313 AA.
 AC Q9QXV9;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE SPROUTY HOMOLOG 1 (SPRY-1).
 GN SPRY1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=99429807; PubMed=10498682;
 RA Minowada G., Jarvis L.A., Chi C.L., Neubuser A., Sun X., Hachon N.,
 RA Krasnow M.A., Martin G.R.;
 RT "Vertebrate sprouty genes are induced by FGF signaling and can cause
 RT chondrodysplasia when overexpressed.";
 RL Development 126:4465-4475(1999).
 CC -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
 CC FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
 CC ORGANOGENESIS.
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS
 CC BUT IS TRANSLOCATED TO THE MEMBRANE RUFFLES IN CELLS STIMULATED
 CC WITH EGF (EPIDERMAL GROWTH FACTOR).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYO AND ADULT TISSUES
 CC INCLUDING HEART, BRAIN, LUNG, KIDNEY, AND SKELETAL MUSCLE.
 CC -1- DEVELOPMENTAL STAGE: AT E8.5 EXPRESSED IN THE PRIMITIVE STREAK,
 CC ROSTRAL FOREBRAIN, CELLS LATERAL TO THE POSTERIOR HINDRAIN,
 CC ANTERIOR HINDRAIN AND DEVELOPING MIDRAIN. AT E9.5 CONTINUES TO
 CC BE EXPRESSED IN THE ROSTRAL FOREBRAIN AND PRIMITIVE STREAK, AND IS
 CC ALSO DETECTED IN THE BRANCHIAL ARCHES AND THE FORELIMB BUD. AT
 CC E10.5 EXPRESSED IN THE SOMITES, FRONTONASAL PROCESSES, TAILBUD,
 CC AND HINDLIMB BUD.
 CC -1- INDUCTION: BY FGF SIGNALING.
 CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
 CC THE PROTEIN TO THE MEMBRANE RUFFLES.
 CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF176903; AAD56004.1; -
 CC Developmental protein: Membrane.
 CC KW MGD; MG11345139; SPY1.
 CC FT DOMAIN 116 125 POLY-SER.
 CC SEQUENCE 313 AA; 34004 MW; CD3C27F7022C0B99 CRC64;

Query Match 40.3%; Score 667; DB 1; Length 313;
 Best Local Similarity 44.6%; Pred. No. 1.8e-43;
 Matches 135; Conservative 46; Mismatches 78; Indels 44; Gaps 8;

QY 16 VMVQPLT---DSMSHSRLQHPPLTLPIDQVKTSHVENDYIDNPSLALTTGPKRTGRGAP 72
 DB 14 VVIQPAVEGRQRLDYDRDTPATILSLDQIRAKINGSNEYTEGSAV-----RRPAP 65
 QY 73 ELATPARCDDVTHNT-----SPSGRP-----SSVSSSSST 105
 DB 66 RTAPRPER--QERTHEIPANVNSSEYHRPASHPNANGSVLSRSTGSAASSGSSSV 123
 QY 106 SSDRLIDHMAPP-PVADQASPAVRIQPKVYHCOPLDKGPAPVELDKHFLICEACKG 164
 DB 124 SSEQGLGRSPTRPIPIGRHSRDRVIRTKQLIVE--DLKA-SLKEDPTGKFKCEQCK 180
 QY 165 CKKCKEASPTLLPSCWGNQECLSAQTLVNYGTGCMQVGIYFHCNTNDEGSCADHRC 224
 DB 181 CKGECTAPRALPSCALCKQCLCSAESNVEYGTGCMQVGIYFHCNDDGGSGYSINRP 240
 QY 225 SCRSNCCARMSFGALSVVLPCLLYLPATGCVKLAQRGYDLRRPGCRCKHTNYSICK 284
 DB 241 SCSQSHCSRYLCMGALSLCLPCLLYCPYPAKCKLTCGXYDMTHRPGCRCKRNSNTYCK 300
 QY 285 AAS 287
 DB 301 LES 303

RESULT 5

SQ SEQUENCE 315 AA: 34623 MW: 81514698EAD809A7 CRC64:
 Query Match 38.2%; Score 633.5; DB 1; Length 315;
 Best Local Similarity 43.3%; Pred. No. 5.9e-41;
 Matches 129; Conservative 46; Mismatches 84; Indels 39; Gaps 9;
 QY 19 QPLIDVSHSHR-----LQPLILPILPDKVTSVENDYIDNPISLALTYGPK--- 65
 DB 12 QPLIDVSHSHR-----LQPLILPILPDKVTSVENDYIDNPISLALTYGPK--- 71
 QY 66 -----RTG-----CAPLAPT-----PARCQDVTTHMISFGSPSSVSSSTSSD 108
 DB 72 RPTQKHEHRLHGLPBNHROPRLQPSQVHSSRRLSNSISTVSSGSSSTRTSTSSSSE 131
 QY 109 QRL--DHMAPPVDAQSPRAVRIOPKVHCOPLDKGPAVPELDKHFLLCEAGCKCK 166
 DB 132 QRLGFSFSGHGAAD-----GIIRVQPK-SELKPGDVK-PLSKDQDLHAYRCEDCGCK 185
 QY 167 CKRCASPRTLPSQWONOCESQVOTLVNCTGCMCLVQGIFFHCTNDEGSCADHPSC 226
 DB 186 CKRCYTRPLPSWICDKOCQCSAONVIDYGTVCYKGLFTHCSN-DDEDNCADNPSC 244
 QY 227 SRSNCCARNSFGALSVLPCLLCYLPATGCVKLAQRYDLRRPGCRCKHTSVIC 284
 DB 245 SSSHCTRMSAGVNSLFLPCLMICYLPKAGCLKLCQCYDRVNRPGCRCKNSYVCK 302
 RESULT 7
 SPY3_HUMAN ID SPY3_HUMAN STANDARD: PRT: 288 AA.
 AC 043610;
 DT 20-AUG-2001 (rel. 40, Created)
 DT 20-AUG-2001 (rel. 40, Last sequence update)
 DE 20-AUG-2001 (rel. 40, Last annotation update)
 GN SPRTY3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RA MEDLINE-2012249; PubMed-10655549;
 RA Cicciocioppa A., D'Esposito M., Esposito T., Gianfrancesco F.,
 RA Migliaccio C., Miano M.G., Matarazzo M.R., Vacca M., Pranze A.,
 RA Cucchiare M., Cocchia M., Curci A., Terracciano A., Torino A.,
 RA Schlessinger D., D'Urso M.;
 RA "Differentially regulated and evolved genes in the fully sequenced
 RA Xq/qy pseudautosomal region.";
 RT Hum. Mol. Genet. 9:395-401(2000).
 RL (2)
 RP SEQUENCE OF 92-191 FROM N.A.
 RX MEDLINE-98117253; PubMed-9458049;
 RA Hachoen N., Kramer S., Sutherland D., Hiromi Y., Krasnow M.A.;
 RT "Sprouty encodes a novel antagonist of FGF signaling that patterns
 RT apical branching of the Drosophila airways.";
 RL Cell 92:253-263(1998).
 CC -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
 CC FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
 CC ORGANOGENESIS.
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS
 CC BUT IS TRANSLOCATED TO THE MEMBRANE RUFLIES IN CELLS STIMULATED
 CC WITH EGF (EPIDERMAL GROWTH FACTOR).
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED; PARTICULARLY IN THE FETAL
 CC TISSUES.
 CC -1- INDUCTION: BY FGF SIGNALING.
 CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
 CC THE PROTEIN TO THE MEMBRANE RUFLIES.
 CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ27133; CAB96768.1; -
 DR EMBL: AF041038; AAC39567.1; -
 DR MIM: 602467; -
 DR InterPro: IPR003019; Metallothion.
 DR Pfam: PF00131; metalthio; 1.
 KW Developmental protein; Membrane.
 FT Domain 147 270
 FT CONFLICT 92 95
 FT TRAS -> PPLP (IN REF. 2).
 SQ SEQUENCE 288 AA: 31221 MW: 620595F5E897D99B8 CRC64:
 Query Match 37.6%; Score 622.5; DB 1; Length 288;
 Best Local Similarity 41.6%; Pred. No. 3.6e-40;
 Matches 123; Conservative 40; Mismatches 62; Indels 71; Gaps 9;
 QY 37 ILPDKVTSVENDYIDNP-----SLA-LTTPKRTGCAPE 73
 DB 12 ILPDKVTSVENDYIDNP-----SLA-LTTPKRTGCAPE 71
 QY 74 LAPPARCDVTHMISFGSPSSVSS---STSSDQDLDMAPPVDAQSPRAVR 130
 DB 72 LQPLPOHLSQ-----SSIASMSHSTTASDORLSTYSPSP---SQQSTIR 114
 QY 131 IOPKVVHCOPLDKGPAVPELD-----KHFLCEAGCKCKCECASPT 175
 DB 115 TOP-----GAGVHPRADALKGFAEQSAGHSEHFLFCEGCRCKPCPTAAR 163
 QY 176 LPSQWONOCESQVOTLVNCTGCMCLVQGIFFHCTNDEGSCADHPSCRSNCARW 235
 DB 164 LPSQWONOCESQVOTLVNCTGCMCLVQGIFFHCTNDEGSCADHPSCRSNCARW 222
 QY 236 SFGALSVLPCLLCYLPATGCVKLAQRYDLRRPGCRCK-HTNSVYICKAASGDA 290
 DB 223 AAMSLISLFLPCLLCYLPATGCVKLAQRYDLRRPGCRCKHTNSVYICKAASGDA 278
 RESULT 8
 SPY1_HUMAN ID SPY1_HUMAN STANDARD: PRT: 139 AA.
 AC 043609;
 DT 20-AUG-2001 (rel. 40, Created)
 DT 20-AUG-2001 (rel. 40, Last sequence update)
 DE 20-AUG-2001 (rel. 40, Last annotation update)
 GN SPROUTY.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98117253; PubMed-9458049;
 RA Hachoen N., Kramer S., Sutherland D., Hiromi Y., Krasnow M.A.;
 RT "Sprouty encodes a novel antagonist of FGF signaling that patterns
 RT apical branching of the Drosophila airways.";
 RL Cell 92:253-263(1998).
 CC -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
 CC FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
 CC ORGANOGENESIS.
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS
 CC BUT IS TRANSLOCATED TO THE MEMBRANE RUFLIES IN CELLS STIMULATED
 CC WITH EGF (EPIDERMAL GROWTH FACTOR).
 CC -1- INDUCTION: BY FGF SIGNALING.
 CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
 CC THE PROTEIN TO THE MEMBRANE RUFLIES.
 CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF041037; AAC39566.1; -
DR MIM: 602465; -
KM Developmental protein; Membrane.
FT NON_TER 1 130 CYS-RICH.
FT DOMAIN 1 130
SQ SEQUENCE 139 AA; 15359 MW; BEC8DAE42B322220 CRC64;

Query Match 31.9%; Score 528; DB 1; Length 139;
Best Local Similarity 62.6%; Pred. No. 2,3e-33;
Matches 87; Conservative 24; Mismatches 24; Indels 4; Gaps 1;

OY 159 CACGCKCKCKECASPTLPSCWVNCDECLVNYGTGCTVGVGIFHCTNDEDEGS 218
DB 1 CACGCKCKCKGCTAPRTLPSCACNRCCLCSASVWEYGTGCTVGVGIFHCSNDEGDS 60
OY 219 GADHPCSCSBNCCARSPFGALSYVLPCLCTLPATGCVKLAQRCYRLRPGCRCKHT 278
DB 61 YSDNCCSCSQSHCCSRVLCMGANSLFLPCLCTPPAKGLCKLRRCYDMIRPGCRKNS 120
OY 279 NSVICKAAS---GDANTS 293
DB 121 NTVCCKLESPPSCGCKPS 139

RESULT 9
SPY_DROME STANDARD: PRT; 589 AA.
AC 044783; Q9VZP7;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROTEIN SPROUTY (SPRY).
GN STY OR CG1921.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-CANTON-S;
RX MEDLINE=98117253; PubMed=9458049;
RA Hachem N., Kramer S., Sutherland D., Hironi Y., Krasnow M.A.;
RT "Sprouty encodes a novel antagonist of FGF signalling that patterns
RT apical branching of the Drosophila airways.";
RL Cell 92:253-263(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.D.,
RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brodtker P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

```

```

RA Dodson K., Doup L.E., Downes M., Dusan-Rocha S., Dunkov B.C., Dunn P.,
RA Dublin K.J., Evangelista C.C., Ferrar C., Ferriere S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gang N.S., Gelbart W.M., Glasser K.,
RA Glader A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshtina N.V., Mobarry C., Morris J., Moshtrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Purl V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yeo Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[3]
RP FUNCTION
RX MEDLINE=99387981; PubMed=10457022;
RA Reich A., Sapir A., Shilo B.-Z.;
RT "Sprouty is a general inhibitor of receptor tyrosine kinase
RT signalling.";
RL Development 126:4139-4147(1999).
[4]
RP FUNCTION
RX MEDLINE=99244704; PubMed=10226010;
RA Kramer S., Okabe M., Hachem N., Krasnow M.A., Hironi Y.;
RT "Sprouty: a common antagonist of FGF and EGF signalling pathways in
RT Drosophila.";
RL Development 126:2515-2525(1999).
[5]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLINE=99189751; PubMed=10089881;
RA Casci T., Vinos J., Freeman M.;
RT "Sprouty, an intracellular inhibitor of Ras signalling.";
RL Cell 96:655-665(1999).
-1- FUNCTION: INHIBITOR OF TRACHEAL BRANCHING THAT RESTRICTS BRANCH
BUDDING BY ANTAGONIZING THE BNL-FGF PATHWAY (BNL: BRANCHLESS, AN
FGF INDUCER OF BRANCHING). ACTS AS AN ANTAGONIST OF EGFR-MEDIATED
SIGNALING IN THE EYE (WHERE IT IS IMPORTANT FOR CELL
DETERMINATION) MIDLINE GLIA, CHORDONAL ORGANS, WING AND OVARIAN
FOLLICLE CELLS.
-1- SUBUNIT: INTERACTS WITH DRK AND GAP1 PROTEIN OF THE RAS PATHWAY.
-1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE INNER SURFACE OF THE
PLASMA MEMBRANE.
-1- TISSUE SPECIFICITY: EXPRESSED IN THE EMBRYONIC TRACHEAL SYSTEM,
DEVELOPING EYE IMAGINAL DISK, EMBRYONIC CHORDONAL ORGAN
PRECURSORS, MIDLINE GLIA, WING DISK AND OVARY.
-1- DEVELOPMENTAL STAGE: FROM STAGE 7 OF OOGENESIS FOUND IN THE
POSTERIOR FOLLICLE CELLS AND DURING STAGE 9 WHEN THE FOLLICLE
CELLS MIGRATE POSTERIORLY OVER THE OOCYTE NUCLEUS, EXPRESSION IS
SEEN IN THE DORSAL AND LATERAL CELLS, AND IS EXCLUDED FROM THE
VENTRAL CELLS. ONCE THE MIGRATION OF FOLLICLE CELLS IS COMPLETE
EXPRESSED IN THE DORSAL-ANTERIOR CORNER OF THE EGG CHAMBER.
-1- INDUCTION: BY THE BNL-FGF PATHWAY IN THE TRACHEAL SYSTEM AND BY
THE EGF RECEPTOR PATHWAY IN THE WING IMAGINAL DISK AND THE
FOLLICLE CELLS OF THE OVARY.
-1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
THE PROTEIN TO THE PLASMA MEMBRANE.
-1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```


CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AR039842; AAC04257.1; -
 DR EMBL: AE003478; AAF4772.1; -
 DR FlyBase: Fggn0014388; scy.
 KW Developmental protein; Membrane.
 FT DOMAIN 104 135 SER-RICH.
 FT DOMAIN 148 162 ASN-RICH.
 FT DOMAIN 207 271 GLN-RICH.
 FT DOMAIN 378 501 CYS-RICH.
 FT DOMAIN 492 551 GLY-RICH.
 FT DOMAIN 249 262 POLY-GLN.
 FT DOMAIN 514 514 POLY-GLY.
 FT CONFLICT 38 38 N -> T (IN REF. 1).
 FT CONFLICT 262 262 Q -> Q00 (IN REF. 1).
 FT CONFLICT 312 312 P -> L (IN REF. 1).
 SQ SEQUENCE 589 AA; 62388 MW; 7450E11987B17B20 CRC64;
 Query Match 23.6%; Score 391; DB 1; Length 589;
 Best Local Similarity 29.3%; Pred. No. 1,9e-22;
 Matches 95; Conservative 51; Mismatches 116; Indels 62; Gaps 11;
 QY 4 PIPQAPLTPNSVYVQPLDSSNMSH-SRLQHPITLPIQDKVTSHEVDYIDNPISALT 62
 DB 190 PRESELTNEYV-----DPLQHATRSQPAQGOONGQTTTHLLLPQRQHHLIQ 243
 QY 63 GPKRTGGAPELAPPARCDQDYTHMISFSGRSSVSSSSSTSSD-----QRLIDM 115
 DB 244 HQCHLD--000000000000HLDHQNOQHARLATVTOATVSGSDHTDGLSHLQNST 301
 QY 116 APPVADQSPRAVRI-----QPKVHCOPLDLK----- 144
 DB 302 TKPPASKOPAPRPLGKGLGLGLGLGNP-IITKQTPPAQKEMHLELQPGGAGN 360
 QY 145 -GPAY-----PELIDKHFLCEAGCKCKEASPRTLPSQVNCQOECTSAOILVYGC 199
 DB 361 GGPLVMAQDPSLLNP--IYCPKRCRCREOCQSPRLPOTVWCNKTKLCSAESVIDYASC 418
 QY 200 MCLVQGIFFHCTNED-----DEGS---CADHPCSCSRNSCCARMSFGALSIVPLCLCYL 252
 DB 419 LCCAKALFHCARDNDLDDGNGTPCVDNPGSCGPKYKRTQRMGIGALSTFLPCLMFYW 478
 QY 253 PATGCYKLAQGYDLRRPGCRCK 276
 DB 479 PMRGCKKICEKCYGRAGRCRCQ 502
 RESULT 10
 SPY1_CHICK STANDARD; PRT; 76 AA.
 ID SPY1_CHICK
 AC O9PTL1;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE SPROUTY HOMOLOG 1 (SPRY-1) (FRAGMENT).
 GN SPRY1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RX MEDLINE=99429807; PubMed=10498682;
 RA Minowada G., Jarvis L.A., Chl C.L., Neubuser A., Sun X., Hachon N.,
 RA Krasnow M.A., Martin G.R.;
 RT "Vertebrate sprouty genes are induced by FGF signaling and can cause
 RT chondrodysplasia when overexpressed.";
 RL Development 126:4465-4475(1999).

CC -1- FUNCTION: MAY FUNCTION AS AN ANTAGONIST OF FIBROBLAST GROWTH
 CC FACTOR (FGF) PATHWAYS AND MAY NEGATIVELY MODULATE RESPIRATORY
 CC ORGANORENESIS.
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE CYTOPLASM IN UNSTIMULATED CELLS
 CC BUT IS TRANSLOCATED TO THE MEMBRANE RUPTLES IN CELLS STIMULATED
 CC WITH EGF (EPIDERMAL GROWTH FACTOR).
 CC -1- TISSUE SPECIFICITY: BRAIN AND INTERLIMB REGION.
 CC -1- INDUCTION: BY FGF SIGNALING.
 CC -1- DOMAIN: THE CYS-RICH DOMAIN IS RESPONSIBLE FOR THE LOCALIZATION OF
 CC THE PROTEIN TO THE MEMBRANE RUPTLES.
 CC -1- DISEASE: OVEREXPRESSION DURING SKELETAL DEVELOPMENT CAUSES A
 CC SEVERE CHONDRODYSPLASIA IN WHICH CARTILAGE CELL DIFFERENTIATION IS
 CC INHIBITED.
 CC -1- SIMILARITY: BELONGS TO THE SPROUTY FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF177875; AAD56008.1; -
 KW Developmental protein; Membrane.
 FT NON_TER 1 67 CYS-RICH.
 FT DOMAIN 1 67
 SQ SEQUENCE 76 AA; 8349 MW; 6B141142FC24C2F8 CRC64;
 Query Match 15.1%; Score 250; DB 1; Length 76;
 Best Local Similarity 57.9%; Pred. No. 1e-12;
 Matches 44; Conservative 13; Mismatches 15; Indels 4; Gaps 1;
 QY 222 HPCSCGRNCCARMSFGALSIVPLCLCYLTPATGCVKLAQRKYDLRRPGCRCKHTSY 281
 DB 1 NPSCGSHSCRCRYLCKAMSLPFLCLCTPPRAGKCLKRCGYDVRNPGCRKNSMTV 60
 QY 282 ICAAS---GDAKTS 293
 DB 61 YCKLESCPSRGOKPS 76
 RESULT 11
 ZAN_PIG STANDARD; PRT; 2476 AA.
 ID ZAN_PIG
 AC Q28583;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE ZONADHESIN PRECURSOR.
 GN ZAN.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 NCBI_TaxID=9823;
 RX MEDLINE=9604658; PubMed=7592795;
 RA Hardy D.M., Garbers D.L.;
 RT "A sperm membrane protein that binds in a species-specific manner to
 RT the egg extracellular matrix is homologous to von Willebrand
 RT factor.";
 RL J. Biol. Chem. 270:26025-26028(1995).
 CC -1- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
 CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
 CC SIGNALING.
 CC -1- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXCLUSIVELY ON THE
 CC APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: IN TESTIS, PRIMARILY IN HAPLOID SPERMATIDS.
 CC NOT IN LUNG, LIVER, HEART, SPLEEN, BRAIN, KIDNEY, EPIDIDYMIS.


```

CC      SIMILARITY).
CC      -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA PAT-3
CC      ASSOCIATES WITH EITHER ALPHA PAT-1 OR ALPHA PAT-2.
CC      -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: U19744; AA85704.1; -
DR      EMBL: Z35604; CAA8467.1; -
DR      WormPep: ZK1058.2; CE01102.
DR      InterPro: IPR000561; EGF-like.
DR      InterPro: IPR001169; Integrin_B.
DR      InterPro: IPR003659; Integrin_beta_C.
DR      InterPro: IPR002035; VWF.
DR      Pfam: PF00362; Integrin_B.1.
DR      PRINTS: PR01186; INTEGRINB.
DR      ProDom: PD00181; Integrin_B.1.
DR      SMART: SM00181; EGF.1.
DR      SMART: SM00187; INB.1.
DR      SMART: SM00423; PSI.1.
DR      SMART: SM00327; VMA.1.
DR      PROSITE: PS00243; INTEGRIN_BETA.2.
DR      PROSITE: PS00022; EGF_1; UNKNOWN.2.
DR      PROSITE: PS01186; EGF_2; UNKNOWN.1.
KW      Integrin, Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW      Signal; Extracellular matrix; Cytoskeleton.
FT      SIGNAL 1..19 POTENTIAL.
FT      CHAIN 20..809 INTEGRIN BETA PAT-3.
FT      DOMAIN 20..737 EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM 738..758 POTENTIAL.
FT      DOMAIN 759..809 CYTOPLASMIC (POTENTIAL).
FT      CARBOHYD 47..47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 141..141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 269..269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 373..373 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 400..400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 530..530 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 672..672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 693..693 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 721..721 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE 809 AA: 90137 MW: 700440189.189 CRC64;
Query Match 6.8%, Score 113; DB 1; Length 809;
Best Local Similarity 26.7%, Pred. No. 0.21;
Matches 27; Conservative 11; Mismatches 37; Indels 26; Gaps 5;
OY 162 CGCCKK-----ECASPTLPS-----CWCNDECLCSAQTIVNYTCMC----- 201
DB 496 CGVCRCKGKGVKCYECNRPMSSTAALNCKRTNENALICEGRGVCNCRGCDNPRANPE 555
OY 202 -LYOGIFYHCTN---EDDEGSCADH-PCSCSRNCARIS 236
DB 556 EOLSGEFCEDCNFCPRHDKRICAEHGECKGCKICACPGMT 596

```

```

OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxId:9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-90062044; PubMed-2584182;
RA      Mancuso D.J., Tuley E.A., Westfield L.A., Morrill N.K.,
RA      Shelton-Inloes B.B., Sorace J.M., Alevy Y.G., Sadler J.E.;
RT      "Structure of the gene for human von Willebrand factor.";
RL      J. Biol. Chem. 264:19514-19527(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-87016349; PubMed-3489923;
RA      Bonthron D., Orr E.C., Mitsuoka L.M., Ginsburg D., Handin R.I.,
RA      Ornkin S.H.;
RT      "Nucleotide sequence of pre-pro-von Willebrand factor cDNA.";
RL      Nucleic Acids Res. 14:7125-7128(1986).
RN      [3]
RP      SEQUENCE OF 1-120 FROM N.A., AND SEQUENCE OF 23-56.
RX      TISSUE-cumbilical vein endothelial cells;
RX      MEDLINE-87213253; PubMed-3495266;
RA      Shelton-Inloes B.B., Broze G.J. Jr., Miletich J.P., Sadler J.E.;
RT      Evolution of human von Willebrand factor: cDNA sequence
RT      polymorphisms, repeated domains, and relationship to von Willebrand
RT      antigen II.";
RL      Biochem. Biophys. Res. Commun. 144:657-665(1987).
RN      [4]
RP      SEQUENCE OF 1-1400 FROM N.A.
RX      MEDLINE-87004550; PubMed-3019665;
RA      Verweij C.L., Diergaarde P.J., Hart M., Pannekoek H.;
RT      "Full-length von Willebrand factor (vWF) cDNA encodes a highly
RT      repetitive protein considerably larger than the mature vWF subunit.";
RL      EMBO J. 5:1839-1847(1986).
RN      [5]
RP      ERRATUM.
RA      Verweij C.L., Diergaarde P.J., Hart M., Pannekoek H.;
RL      EMBO J. 5:3074-3074(1986).
RN      [6]
RP      SEQUENCE OF 764-2813.
RX      MEDLINE-86269895; PubMed-3524673;
RA      Titani K., Kumar S., Takio K., Ericsson L.H., Wade R.D., Ashida K.,
RA      Walsh K.A., Chopok M.W., Sadler J.E., Fujikawa K.;
RT      "Amino acid sequence of human von Willebrand factor.";
RL      Biochemistry 25:3171-3184(1986).
RN      [7]
RP      SEQUENCE OF 781-1424 FROM N.A.
RX      MEDLINE-86269894; PubMed-3488076;
RA      Shelton-Inloes B.B., Titani K., Sadler J.E.;
RT      "cDNA sequences for human von Willebrand factor reveal five types of
RT      repeated domains and five possible protein sequence polymorphisms.";
RL      Biochemistry 25:3164-3171(1986).
RN      [8]
RP      SEQUENCE OF 764-873 AND 1289-2813 FROM N.A.
RX      MEDLINE-86016708; PubMed-2864688;
RA      Sadler J.E., Shelton-Inloes B.B., Sorace J.M., Harlan J.M.,
RA      Titani K., Davie E.W.;
RT      "Cloning and characterization of two cDNAs coding for human von
RT      Willebrand factor.";
RL      Proc. Natl. Acad. Sci. U.S.A. 82:6394-6398(1985).
RN      [9]
RP      SEQUENCE OF 990-1947 FROM N.A.
RX      MEDLINE-91105089; PubMed-1988024;
RA      Mancuso D.J., Tuley E.A., Westfield L.A., Lester-Mancuso T.L.,
RA      Le Beau M.M., Sorace J.M., Sadler J.E.;
RT      "Human von Willebrand factor gene and pseudogene: structural analysis
RT      and differentiation by polymerase chain reaction.";
RL      Biochemistry 30:253-269(1991).
RN      [10]
RP      SEQUENCE OF 2731-2813 FROM N.A.
RX      MEDLINE-85269603; PubMed-3875078;
RA      Verweij C.L., de Vries C.J.M., Distel B., van Zonneveld A.-J.,
RA      Geurts van Kessel A., van Mourik J.A., Pannekoek H.;

```

RT *Construction of cDNA coding for human von Willebrand factor using
RT antibody probes for colony-screening and mapping of the chromosomal
RT gene.";
RL Nucleic Acids Res. 13:4699-4717(1985).
RN (111)
RP SEQUENCE OF 1-177 FROM N.A.
RX MEDLINE-88111704; PubMed-2828057;
RA Bonhron D., Orkin S.H.;
RT "The human von Willebrand factor gene. Structure of the 5' region.";
RL Eur. J. Biochem. 171:51-57(1988).
RN (112)
RP SEQUENCE OF 2621-2813 FROM N.A.
RX MEDLINE-85244588; PubMed-3874428;
RA Ginsburg D., Handlin R.I., Bonthon D.T., Donlon T.A., Bruns G.A.P.,
RA Latt S.A., Orkin S.H.;
RT "Human von Willebrand factor (vWF): isolation of complementary DNA
RT (cDNA) clones and chromosomal localization.";
RL Science 228:1401-1406(1985).
RN (113)
RP SEQUENCE OF 2731-2813 FROM N.A.
RX MEDLINE-85201687; PubMed-3873280;
RA Lynch D.C., Zimmerman T.S., Collins C.J., Brown M., Morin M.J.,
RA Ling E.H., Livingston D.M.;
RT "Molecular cloning of cDNA for human von Willebrand factor:
RT authentication by a new method.";
RL Cell 41:49-56(1985).
RN (114)
RP REVISIONS.
RA Lynch D.C.;
RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
RN (115)
RP SEQUENCE OF 2731-2813 FROM N.A.
RX MEDLINE-87260814; PubMed-3496594;
RA Collins C.J., Underdahl J.P., Levene R.B., Ravera C.P.,
RA Morin M.J., Dombalagian M.J., Ricca G., Livingston D.M.,
RA Lynch D.C.;
RT "Molecular cloning of the human gene for von Willebrand factor and
RT identification of the transcription initiation site.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4393-4397(1987).
RN (116)
RP DISULFIDE BONDS.
RX MEDLINE-88163465; PubMed-3502076;
RA Marti T., Roselet S.J., Titani K., Walsh K.A.;
RT "Identification of disulfide-bridged substructures within human von
RT Willebrand factor.";
RL Biochemistry 26:8099-8109(1987).
RN (117)
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE-86274702; PubMed-3089784;
RA Samor B., Michalski J.C., Debray H., Mazurier C., Goudemand M.,
RA van Halbeek H., Vliegenhart J.F.G., Montreuil J.;
RT "Primary structure of a new tetraantennary glycan of the N-
RT acetylglucosaminic type isolated from human factor VIII/von
RT Willebrand factor.";
RL Eur. J. Biochem. 158:295-298(1986).
RN (118)
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 1261-1468.
RX MEDLINE-98221174; PubMed-9553097;
RA Emsley J., Cruz M., Handlin R., Liddington R.;
RT "Crystal structure of the von Willebrand factor A1 domain and
RT implications for the binding of platelet glycoprotein Ib.";
RL J. Biol. Chem. 273:10396-10401(1998).
RN (119)
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1665-1873.
RX MEDLINE-97472939; PubMed-9331419;
RA Huisinga E.G., Martijn van der Plas R., Kroon J., Sixma J.J., Gros P.;
RT "Crystal structure of the A3 domain of human von Willebrand factor:
RT implications for collagen binding.";
RL Structure 5:1147-1156(1997).
RN (120)
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1686-1872.
RX MEDLINE-97460108; PubMed-9312128;
RA Bienkowska J., Cruz M., Altiero A., Handlin R., Liddington R.;

RT "The von Willebrand factor A3 domain does not contain a metal ion-
RT dependent adhesion site motif.";
RL J. Biol. Chem. 272:25162-25167(1997).
RN (121)
RP VARIANTS TRP-1597 AND ASP-1607.
RX MEDLINE-89264495; PubMed-2786201;
RA Ginsburg D., Koukle B.A., Gill J.C., Montgomery R.R.,
RA Bockensiedt P.L., Johnson T.A., Yang A.Y.;
RT "Molecular basis of human von Willebrand disease: analysis of
RT platelet von Willebrand factor mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:3723-3727(1989).
RN (122)
RP VARIANT THR-1628.
RX MEDLINE-91196734; PubMed-1673047;
RA Iannuzzi M.C., Hidaka N., Boehnke M., Bruck M.E., Hanna W.T.,
RA Collins F.S., Ginsburg D.;
RT "Analysis of the relationship of von Willebrand disease (VWD) and
RT hereditary hemorrhagic telangiectasia and identification of a
RT potential type IIA VWD mutation (I1685 to Thr).";
RL Am. J. Hum. Genet. 48:757-763(1991).
RN (123)
RP VARIANTS NORMANDY-2 AND NORMANDY-3.
RX MEDLINE-92001464; PubMed-1832934;
RA Gaucher C., Mercier B., Jorleux S., Oufkir D., Mazurier C.;
RT "Identification of two point mutations in the von Willebrand
RT gene of three families with the 'Normandy' variant of von Willebrand
RT disease.";
RL Br. J. Haematol. 78:506-514(1991).
RN (124)
RP VARIANT CYS-1308.
RX MEDLINE-92104315; PubMed-1761120;
RA Donner M., Andersson A.-M., Kristoffersson A.-C., Nilsson I.M.,
RA Dahlback B., Holmberg L.;
RT "An Arg545--Cys545 substitution mutation of the von Willebrand
RT factor in type IIB von Willebrand's disease.";
RL Eur. J. Haematol. 47:342-345(1991).
RN (125)
RP VARIANTS TRP-1306; CYS-1308 AND PRO-1613.
RX MEDLINE-91185601; PubMed-2010538;
RA Randi A.M., Rabinowitz I., Mancuso D.J., Mannucci P.M., Sadler J.E.;
RT "Molecular basis of von Willebrand disease type IIB. Candidate
RT mutations cluster in one disulfide loop between proposed platelet
RT glycoprotein Ib binding sequences.";
RL J. Clin. Invest. 87:1220-1226(1991).
RN (126)
RP VARIANTS TRP-1306; CYS-1308; MET-1316; GLN-1341 AND HIS-1399.
RX MEDLINE-91185602; PubMed-1672694;
Query Match 6.7%; Score 110.5; DB 1; Length 2813;
Best local similarity 21.0%; Pred. No. 1.1;
Matches 56; Conservative 30; Mismatches 90; Indels 91; Gaps 13;
QY 77 TPARDODVTHHWISFGSRPSSSSSTSPQRLDHWAPPPVADQSPPAVRIQKVV 136
DB 721 TWICEDGDFMH--CTMGVPGSLLPDAVLSSP---LSH-----SKRSLSGRPPMV 772
QY 137 H-----CQPLDK-----GPAVPELDKHFLLCEAGCKKC--- 167
DB 773 KLYCPADNLRAELECTKTKCONYDLCEMSGCVSGCLCPGVAVRHENRCVLEKRC PFHQ 832
QY 168 -KECASPTLLESCVWQNECLCSAQTLVNVGTCMCLVQGFYHCHNEDESSCADHPC-- 224
DB 833 GKRYAPGTTVA-----ICGNTVC-----RRKKNKCDIHYCDA 865
QY 225 SCGRNSCCARMSFGALSVVLPCLLYL-----PATGCVKLAORGDLRRPCCRC 275
DB 866 TCSTGMAHYTFEDG-LKYLFPCECQYVLVDYCGSNPOTFRILVGNKG--CSHPSVKC 921
QY 276 KHTNSVICKAK-----SGDAKTSRDPK 297
DB 922 KKRVITLVEGGETLFDGEVAVVKKRPMK 948

FT DISULFID 789 799 BY SIMILARITY.
 FT DISULFID 806 817 BY SIMILARITY.
 FT DISULFID 811 826 BY SIMILARITY.
 FT DISULFID 828 837 BY SIMILARITY.
 FT DISULFID 844 855 BY SIMILARITY.
 FT DISULFID 849 890 BY SIMILARITY.
 FT DISULFID 892 901 BY SIMILARITY.
 FT DISULFID 908 919 BY SIMILARITY.
 FT DISULFID 913 928 BY SIMILARITY.
 FT DISULFID 930 939 BY SIMILARITY.
 FT DISULFID 946 957 BY SIMILARITY.
 FT DISULFID 952 966 BY SIMILARITY.
 FT DISULFID 968 977 BY SIMILARITY.
 FT DISULFID 984 995 BY SIMILARITY.
 FT DISULFID 989 1009 BY SIMILARITY.
 FT DISULFID 1011 1020 BY SIMILARITY.
 FT DISULFID 1211 1222 BY SIMILARITY.
 FT DISULFID 1216 1231 BY SIMILARITY.
 FT DISULFID 1233 1242 BY SIMILARITY.
 FT DISULFID 1485 1496 BY SIMILARITY.
 FT DISULFID 1490 1505 BY SIMILARITY.
 FT DISULFID 1507 1516 BY SIMILARITY.
 FT DISULFID 1763 1774 BY SIMILARITY.
 FT DISULFID 1768 1783 BY SIMILARITY.
 FT DISULFID 1785 1794 BY SIMILARITY.
 FT DISULFID 1801 1812 BY SIMILARITY.
 FT DISULFID 1806 1821 BY SIMILARITY.
 FT DISULFID 1823 1832 BY SIMILARITY.
 FT DISULFID 1839 1850 BY SIMILARITY.
 FT DISULFID 1844 1859 BY SIMILARITY.
 FT DISULFID 1861 1870 BY SIMILARITY.
 FT DISULFID 1878 1889 BY SIMILARITY.
 FT DISULFID 1883 1903 BY SIMILARITY.
 FT DISULFID 1905 1914 BY SIMILARITY.
 FT DISULFID 1919 1930 BY SIMILARITY.
 FT DISULFID 1924 1939 BY SIMILARITY.
 FT DISULFID 1941 1950 BY SIMILARITY.
 FT DISULFID 1957 1968 BY SIMILARITY.
 FT DISULFID 1962 1977 BY SIMILARITY.
 FT DISULFID 1979 1988 BY SIMILARITY.
 FT DISULFID 1995 2008 BY SIMILARITY.
 FT DISULFID 2002 2017 BY SIMILARITY.
 FT DISULFID 2019 2028 BY SIMILARITY.
 FT CARBOHYD 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 555 555 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 736 736 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 860 860 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 884 884 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 6.5%: Score 108; DB 1; Length 2139;

Best Local Similarity 25.8%: Pred. No. 1.3;

Matches 63; Conservative 15; Mismatches 82; Indels 84; Gaps 15;

QY 68 RCAGELAPTRA-KCDDP-----VTNHWISFSGRPSVSSSSSTSSDQ-----109
 DB 1658 KWMESMVPPTAFSTDDIOAFOSLATSTOYVLGMPSSRQARSTLSAOGSGOFGKCV 1717
 QY 110 ---RLIDMAPP-PVADQASPRVRIQPKVYHCQPLDLGPAVPELDKHFLLCEACGKC 165
 DB 1718 GEARGLDLPYFMAELIYRTNVSQOKA-----QFLNMTREEG-----CILCROS 1766
 QY 166 KKEE---CASRTILPSCVNC-----NOECICSAQTLVYNTGTGCMCLYOG 205
 DB 1767 DCKNDGQSPSPSDEYAC-TCOPGEFGDGCSTDIDECLENTCL-----NNGTCLINDVAA 1818

QY 206 IF-----YHCTNEDDEGSCADHPCSCSRNC-----CARMSFMALSYLP 246
 DB 1819 FFCQCGKRFEGHCEONIDE--CADQPCN-NGNCTDLINASTYCDCEP-DIMGQCDVLR 1874
 QY 247 CLIC 250
 DB 1875 QMTC 1878

RESULT 15

ID BAR3-CHITE STANDARD; PRT: 1700 AA.
 AC 003376;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE BALBIANI RING PROTEIN 3 PRECURSOR.
 GN BR3.
 OS Chironomus tentans (Midge).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
 OC Chironomidae; Chironomidae; Chironominae; Chironomus.
 OX NCBI_Taxid=7153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Salivary gland;
 RX MEDLINE=90172404; Pubmed=1689777;
 RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
 RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
 repetitive structure split by many introns.";
 RL J. Mol. Biol. 211:331-349(1990).
 CC -FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
 STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
 ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
 INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
 FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
 CC -TISSUE SPECIFICITY: SECRETED.
 CC -DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
 CC -
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X52263; CAA36506.1; -
 CC DR PIR; S08167; S08167.
 CC DR HSSP; P18055; 2MRB.
 CC DR InterPro; IPR000853; Mtl1bion_nemat.
 CC DR PRINTS; PR00876; MTNEMATODE.
 CC Repeat; Signal.
 FT SIGNAL 1
 FT CHAIN 1 1700
 FT SEQUENCE 1700 AA; 186145 MW; 34202828521B0815 CRC64;

Query Match 6.5%: Score 107.5; DB 1; Length 1700;

Best Local Similarity 27.3%: Pred. No. 1.1;

Matches 44; Conservative 17; Mismatches 49; Indels 51; Gaps 14;

QY 138 CQPLDLGPAVPELDKHFLLCEACG-----KCKCEKASPTLPSC-----WVC 182
 DB 49 CYP1---GALFDPK-SCTFMSQSKCSAKQIFNKKACAC-ECRNDPKDCGAGRYW-C 102
 QY 183 NOECICSAQTLVYNTGTGCMCLYOGIFVHCTNEDDEGSCADHPCSCSRNCAR--NSFMG 239
 DB 103 NQDCSCKSTPMSAGCC-----SGSQIWC-----EKSACA---CVPNADKCTAPQVYWKD 150
 QY 240 ALSYVLPCLLCYL-----PATGVKYLQAGRYDRLRRPGCRK 276

Db 151 C-----CCGCPVNMOPADGCTK--PLIMDKV---DCRCE 180

Search completed: May 8, 2002, 10:49:52
Job time: 189 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2002, 02:17:35 ; Search time 2645.75 Seconds

(without alignments)
13590.551 Million cell updates/sec

Title: US-10-082-902-1

Perfect score: 1 agtagcagactgagctgac.....tgtcttgagggaagctgtaga 1662

Sequence: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Scoring table: 21979536 segs, 10817449327 residues

Searched: Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_Main:*

1: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US081.COMB.seq.*

Result				Query				SUMMARIES				Description	
No.	Score	Match	Length	DB	ID								
1	1662	100.0	1662	17	US-09-370-398-1							Sequence 1, Appl	
2	1662	100.0	1662	37	US-10-083-215-1							Sequence 1, Appl	
3	1662	100.0	1662	37	US-10-090-190-1							Sequence 1, Appl	
4	1640	98.7	2061	31	US-09-814-354-2							Sequence 2, Appl	
5	1324.8	79.7	1507	1	PCT-US01-27760A-373							Sequence 373, App	
6	1324.8	79.7	1507	1	PCT-US01-27760A-373							Sequence 373, App	
7	1324.8	79.7	1507	1	PCT-US01-27760A-373							Sequence 373, App	
8	1203.8	72.4	1208	18	US-09-471-275-5770							Sequence 5770, App	
9	889.6	53.5	1402	25	US-09-360-207-6794							Sequence 6794, App	
10	688.8	41.4	901	24	US-09-620-392-36647							Sequence 36647, A	
11	675.8	40.7	680	18	US-09-489-036-34213							Sequence 34213, A	
12	675.8	40.7	680	35	US-09-943-143-34213							Sequence 34213, A	
13	658	39.6	658	18	US-09-457-877-66							Sequence 66, Appl	
14	658	39.6	658	18	US-09-952-981-66							Sequence 66, Appl	
15	583.6	35.1	635	35	US-09-489-036-15011							Sequence 15011, A	
16	583.6	35.1	635	35	US-09-943-143-15011							Sequence 15011, A	
17	580.6	34.9	634	18	US-09-489-036-34210							Sequence 34210, A	
18	580.6	34.9	634	35	US-09-943-143-34210							Sequence 34210, A	
19	569.4	34.3	634	35	US-09-489-036-34209							Sequence 34209, A	
20	569.4	34.3	634	35	US-09-943-143-34209							Sequence 34209, A	
21	561.6	33.8	635	18	US-09-489-036-15010							Sequence 15010, A	
22	561.6	33.8	635	18	US-09-943-143-15010							Sequence 15010, A	
23	476.4	28.7	546	23	US-09-607-200-7267							Sequence 7267, App	
24	476.4	28.7	546	23	US-09-649-163-10408							Sequence 10408, A	
25	466.2	28.1	559	31	US-09-702-134-14899							Sequence 14899, A	
26	466.2	28.1	559	31	US-09-815-264-88947							Sequence 88947, A	
27	443.4	26.7	449	18	US-09-471-275-4064							Sequence 4064, App	
28	443.4	26.7	449	70	US-09-311-261-2444							Sequence 2444, App	
29	421.6	25.4	429	18	US-09-489-036-986							Sequence 986, App	
30	421.6	25.4	429	30	US-09-943-143-986							Sequence 986, App	
31	417	25.1	423	19	US-09-528-409-114858							Sequence 114858, A	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

QY	421	gcccacagagctgccccgcagacgcgcgcgcgtctgacacagatgctcacccacattgatac	480
Db	421	gcccacagagctgccccgcagacgcgcgcgcgtctgacacagatgctcacccacattgatac	480
QY	481	lcccttcagagggcgcccccagatctctgtgagagagagacagacacatcccttcagacacg	540
Db	481	lcccttcagagggcgcccccagatctctgtgagagagagacacatcccttcagacacg	540
QY	541	ctcttcagacacatagtgacacacacaccccgctgctgcgcacagagctcacacacagagctggcgc	600
Db	541	ctcttcagacacatagtgacacacacaccccgctgctgcgcacagagctcacacacagagctggcgc	600
QY	601	atccacagcccaaggtgtgtccacctgcacagacgcctgtgacactcaagagcccggtgtccacac	660
Db	601	atccacagcccaaggtgtgtccacctgcacagacgcctgtgacactcaagagcccggtgtccacac	660
QY	661	gagctctgacacagacattctctgtctgtgtgagagccctctgtgagagtgtaaatgcaagagtgct	720
Db	661	gagctctgacacagacattctctgtctgtgtgagagcccgctgtgtgagagtgtaaatgcaagagtgct	720
QY	721	gcatacccccgcagagctgtgctctccctctgcgtgtcgcgcacacagagagtgctcgtgtcagac	780
Db	721	gcatacccccgcagagctgtgctctccctctgcgtgtcgcgcacacagagagtgctcgtgtcagac	780
QY	781	caaacactctgtacaaatactgagagacgtgtgacatgtctgtgtgagagagacattcttcacacgtc	840
Db	781	caaacactctgtgtacaaatactgagagacgtgtgacatgtctgtgtgagagagacattcttcacacgtc	840
QY	841	acgaaatgagagagatgagatgtcctcctgcgcgtgcacaccccgtctccgtctccgcctcacaac	900
Db	841	acgaaatgagagagatgagatgtcctcctgcgcgtgcacaccccgtctccgtctccgcctcacaac	900
QY	901	tgtctgtgccccgtctgtgtccctctcatctgtgtgcctctccgtgtgtgtgcctcgtctgcgtc	960
Db	901	tgtctgtgccccgtctgtgtccctctcatctgtgtgcctctccgtgtgtgtgcctcgtctgcgtc	960
QY	961	tacacctgcctgcacacagagctgtgtgtgagagctgtgcgcacagagtgagctctgacacgcgtctgcgcgcgc	1020
Db	961	tacacctgcctgcacacagagctgtgtgtgagagctgtgcgcacagagtgagctctgacacgcgtctgcgcgcgc	1020
QY	1021	ccctgtctgtgcctgtgcacagacacagaaacagctgtacatctgtcaaaagacagaccagctggagatgc	1080
Db	1021	ccctgtctgtgcctgtgcacagacacagaaacagctgtacatctgtcaaaagacagaccagctggagatgc	1080
QY	1081	aagacacagacagagcccgagaaagacccctctctgacagcttgtgtgtgagagcccccgtgtgcctgtgc	1140
Db	1081	aagacacagacagagcccgagaaagacccctctctgacagcttgtgtgtgagagcccccgtgtgcctgtgc	1140
QY	1141	tgtgaaacctgtgtctctctctctgcacatctaaagaaagactgtcagacagatctagaagctttttagcc	1200
Db	1141	tgtgaaacctgtgtctctctctctctgcacatctaaagaaagactgtcagacagatctagaagctttttagcc	1200
QY	1201	tccctgtgagctgtgcctctgtctgtctgcgcacatccctcaacccacagcttctctgagaaataacagag	1260
Db	1201	tccctgtgagctgtgcctctgtctgtctgcgcacatccctcaacccacagcttctctgagaaataacagag	1260
QY	1261	accacacacacacgtacacctgtatctccccaagatgatagtaaagaaacatttgtgggctttttt	1320
Db	1261	accacacacacacgtacacctgtatctccccaagatgatagtaaagaaacatttgtgggctttttt	1320
QY	1321	caaggtgtccctgaaactttgtgtctcaaaacagacatgtacagggcgagaggtgtgtgtctgtgggggga	1380
Db	1321	caaggtgtccctgaaactttgtgtctcaaaacagacatgtacagggcgagaggtgtgtgtctgtgggggga	1380
QY	1381	aattttctcttttcagaagacagaaacacagatgtgtgacacataccggaaactgtcagctgt	1440
Db	1381	aattttctcttttcagaagacagaaacacagatgtgtgacacataccggaaactgtcagctgt	1440
QY	1441	ctgtgaatgtccttcccaagccctctctctccctccctccctccctccctccctctctctt	1500
Db	1441	ctgtgaatgtccttcccaagccctctctctccctccctccctccctccctccctctctctt	1500
QY	1501	tccatattcttcttgactcaagagagctagctgtccctgtggaggaattgtactagtaacc	1560

Db 1501 tccatgtcttcttgacacacagagctagctgctggaggaattgtaacagtaacc 1560
Oy 1561 aggttaccttaagaagacccttgagatcttctatactcttctctcccatctca 1620
Db 1561 aggttaccttaagaagacccttgagatcttctatactcttctctcccatctca 1620
Oy 1621 cccacacccacttctgctctgatacttcttggggaagttaga 1662
Db 1621 cccacacccacttctgctctgatacttcttggggaagttaga 1662

RESULT 2

US-10-083-215-1
Sequence 1, Application US/10083215
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
APPLICANT: Montgomery, Julie R.
TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
FILE REFERENCE: 28110/35878
CURRENT FILING DATE: 2002-02-26
PRIOR APPLICATION NUMBER: US/09/370,398
PRIORITY FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1662
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (211)..(1107)
US-10-083-215-1

Query Match 100.0% Score 1662: DB 37: Length 1662:
Best Local Similarity 100.0%: Pred. NO. 0;
Matches 1662: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 aggttaccttaagaagacccttgagatcttctatactcttctctcccatctca 60
Db 1 aggttaccttaagaagacccttgagatcttctatactcttctctcccatctca 60
Oy 61 gttctctccagcggatctgagatgacagcgagagctgagctgagagcgagagcttgcg 120
Db 61 gttctctccagcggatctgagatgacagcgagagctgagctgagagcgagagcttgcg 120
Oy 121 acacagacgttgggagctgagatcttctgagacgtctctcaaaagcccttagaagctgt 180
Db 121 acacagacgttgggagctgagatcttctgagacgtctctcaaaagcccttagaagctgt 180
Oy 181 tttctccgtatagctcagagctcagagcccatctgagcccccagatcccaagagcgcccc 240
Db 181 tttctccgtatagctcagagctcagagcccatctgagcccccagatcccaagagcgcccc 240
Oy 241 ttgactcccaactcagatctgagctcagcccttctgacagcgagatgtctccacacg 300
Db 241 ttgactcccaactcagatctgagctcagcccttctgacagcgagatgtctccacacg 300
Oy 301 ctccagacacccactcactctactacacatgacagagtgaaagacacacatgtgaggaat 360
Db 301 ctccagacacccactcactctactacacatgacagagtgaaagacacacatgtgaggaat 360
Oy 361 gactactatagaacacacttgccttgagccttgacacacgagcccaagcgagacggcgag 420
Db 361 gactactatagaacacacttgccttgagccttgacacacgagcccaagcgagacggcgag 420
Oy 421 gccacagagctggcccgagcccgagcctgagacagagtgatcacaacacacatgtgac 480
Db 421 gccacagagctggcccgagcccgagcctgagacagagtgatcacaacacacatgtgac 480
Oy 481 tccctcagcgggcgccccagctctgtgagcagcagcagcagacata tccctctgacacag 540

Db 481 tccctcagcgggcgccccagctctgtgagcagcagcagcagcagacatcctctgacacag 540
Oy 541 ctctttagacacatgagcaccacacacccgctgagacagcctcacaagagctgtgcg 600
Db 541 ctctttagacacatgagcaccacacacccgctgagacagcctcacaagagctgtgcg 600
Oy 601 atccacacacagtgatctcagctcagcgccttgagacacagggcccgagctccaccc 660
Db 601 atccacacacagtgatctcagctcagcgccttgagacacagggcccgagctccaccc 660
Oy 661 gagctgagacagacactctctgtctgagagcctgtggaagtgtataatgacagagatgt 720
Db 661 gagctgagacagacactctctgtctgagagcctgtggaagtgtataatgacagagatgt 720
Oy 721 gcatcccccagagcttgcctctcagctgagctgagctgagacacagagagctgtcagcc 780
Db 721 gcatcccccagagcttgcctctcagctgagctgagctgagacacagagagctgtcagcc 780
Oy 781 cagactctgtcactaactatgacagctgacatgtgttctgagagagcactctctacacatgc 840
Db 781 cagactctgtcactaactatgacagctgacatgtgttctgagagagcactctctacacatgc 840
Oy 841 acgaatgagagacatgagagctcctgagctgacacacccctgctcctcccgctccac 900
Db 841 acgaatgagagacatgagagctcctgagctgacacacccctgctcctcccgctccac 900
Oy 901 tgcctgcccgcctgtgtctctcaatgagctgtctctcctgagctgtgctcctgctgtgc 960
Db 901 tgcctgcccgcctgtgtctctcaatgagctgtctctcctgagctgtgtcctgctgtgc 960
Oy 961 taccctgctcagcagcctgagctgagagctgagcagctgagcagcagctgagcagc 1020
Db 961 taccctgctcagcagcctgagctgagagctgagcagctgagcagcagctgagcagc 1020
Oy 1021 cctgtgttcgcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
Db 1021 cctgtgttcgcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
Oy 1081 aagacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
Db 1081 aagacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
Oy 1141 tggaaacccgtgtctctctctgacatcactaagaagctgacagagctcagaggttttaagcc 1200
Db 1141 tggaaacccgtgtctctctctgacatcactaagaagctgacagagctcagaggttttaagcc 1200
Oy 1201 tccctgagctgacactgtctgagctgagccacccctacccacagctctcggaaataacagag 1260
Db 1201 tccctgagctgacactgtctgagctgagccacccctacccacagctctcggaaataacagag 1260
Oy 1261 accacacacagctacccgtatctcccaagatgagtaagaagacacttggggcttttt 1320
Db 1261 accacacacagctacccgtatctcccaagatgagtaagaagacacttggggcttttt 1320
Oy 1321 caggtgtccctgaaacttctgtcacaacagacagctcagggcgaggtgtgttgggggga 1380
Db 1321 caggtgtccctgaaacttctgtcacaacagacagctcagggcgaggtgtgttgggggga 1380
Oy 1381 aattttctcttctcagaagacagaaacagatgtgagacacatctccggaactctgagctg 1440
Db 1381 aattttctcttctcagaagacagaaacagatgtgagacacatctccggaactctgagctg 1440
Oy 1441 ctgaaatgctctccagccctctctctcctctccctcctccctccctctctct 1500
Db 1441 ctgaaatgctctccagccctctctctcctctccctcctccctccctctctct 1500
Oy 1501 tccatgtcttctgagacacagagctgagctgagctgagagagagctgacaaatgtgagc 1560
Db 1501 tccatgtcttctgagacacagagctgagctgagctgagagagagctgacaaatgtgagc 1560
Oy 1561 aggttaccttaagaagacccttgagatcttctatactcttctctcccatctca 1620


```
RESULT 4
Sequence 2, Application US/09014354
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Wehrman, Tom
APPLICANT: Yang, Yonhong
APPLICANT: Xue, Aldong J.
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 784CIP2G
CURRENT APPLICATION NUMBER: US/09/814,354
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PL_FL_genes Version 2.0
SEQ ID NO 2
LENGTH: 2061
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (245)..(1141)
US-09-814-354-2

Query Match      98.7% Score 1640; DB 31; Length 2061;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 3 gtagagagctgagctgacagcgagagctgagctgtagagcgagagagccttgcggt 62
Db 37 gtagagagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 96
QY 63 tccctcagcagcgctctgagctgagctgagctgagctgagctgagctgagctgagct 122
Db 97 tccctcagcagcgctctgagctgagctgagctgagctgagctgagctgagctgagct 156
QY 123 acagagcgtaggagctgagctgagctgagctgagctgagctgagctgagctgagct 182
Db 157 acagagcgtaggagctgagctgagctgagctgagctgagctgagctgagctgagct 216
QY 183 ctccgtagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 242
Db 217 ctccgtagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 276
QY 243 gaetcccaactgagctgagctgagctgagctgagctgagctgagctgagctgagct 302
Db 277 gaetcccaactgagctgagctgagctgagctgagctgagctgagctgagctgagct 336
QY 303 ccaatcacccatcacctcacctcacctcacctcacctcacctcacctcacctcacct 362
Db 337 ccagcacccatcacctcacctcacctcacctcacctcacctcacctcacctcacct 396
QY 363 ctacatagacacccatgagctgagctgagctgagctgagctgagctgagctgagct 422
Db 397 ctacatagacacccatgagctgagctgagctgagctgagctgagctgagctgagct 456
QY 423 ccagagctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 482
Db 457 ccagagctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 516
QY 483 ctcaagcgagcgagctgagctgagctgagctgagctgagctgagctgagctgagct 542
|||||
```

```
Db 517 ctcaagcgagcgagctgagctgagctgagctgagctgagctgagctgagctgagct 576
QY 543 cttaagacacatgagctgagctgagctgagctgagctgagctgagctgagctgagct 602
Db 577 cttaagacacatgagctgagctgagctgagctgagctgagctgagctgagctgagct 636
QY 603 ccagcccaagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 662
Db 637 ccagcccaagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 696
QY 663 gctgagacacaccttctgagctgagctgagctgagctgagctgagctgagctgagct 722
Db 697 gctgagacacaccttctgagctgagctgagctgagctgagctgagctgagctgagct 756
QY 723 atcccccgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 782
Db 757 atcccccgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 816
QY 783 gactctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 842
Db 817 gactctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 876
QY 843 gaattgagacagctgagctgagctgagctgagctgagctgagctgagctgagct 902
Db 877 gaattgagacagctgagctgagctgagctgagctgagctgagctgagctgagct 936
QY 903 ctgagcgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 962
Db 937 ctgagcgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 996
QY 963 cctgctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1022
Db 997 cctgctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1056
QY 1023 tggctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1082
Db 1057 tggctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1116
QY 1083 gaccagacagcgagctgagctgagctgagctgagctgagctgagctgagctgagct 1142
Db 1117 gaccagacagcgagctgagctgagctgagctgagctgagctgagctgagctgagct 1176
QY 1143 gaacacctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1202
Db 1177 gaacacctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1236
QY 1203 ctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1262
Db 1237 ctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1296
QY 1263 caccacacgtagctgagctgagctgagctgagctgagctgagctgagctgagct 1322
Db 1297 caccacacgtagctgagctgagctgagctgagctgagctgagctgagctgagct 1356
QY 1323 gggctcctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1382
Db 1357 gggctcctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1416
QY 1383 ttcttcttctcagagacagacacagatgagctgagctgagctgagctgagctgagct 1442
Db 1417 ttcttcttctcagagacagacacagatgagctgagctgagctgagctgagctgagct 1476
QY 1443 tgaatgcttccagcgcttcttcttcttcttcttcttcttcttcttcttcttctt 1502
Db 1477 tgaatgcttccagcgcttcttcttcttcttcttcttcttcttcttcttcttctt 1536
QY 1503 catgtcttctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1562
Db 1537 catgtcttctgagctgagctgagctgagctgagctgagctgagctgagctgagct 1595
QY 1563 gtaaccttaagacacacctgagcttctataaccttcttcttcttcttcttcttct 1622
Db 1596 gtaaccttaagacacacctgagcttctataaccttcttcttcttcttcttcttct 1655
|||||
```


PCT-US01-27760A-373

Query Match 79.7%: Score 1324.8; DB 1; Length 1507;
Best Local Similarity 99.1%; Pred. No. 2.4e-285;
Matches 1332; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy 319 atcctaccatgacaggtgaaagacacacatgttgagaaatgacataagacaacct 378
Db 163 accctgacgtcctcacaagtgaagacagacatgtggaatgacataagacaacct 222
Qy 379 agccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 438
Db 223 agccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 282
Qy 439 agccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 498
Db 283 agccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 342
Qy 499 agccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 558
Db 343 agccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 402
Qy 559 ccacacacacacacacacacacacacacacacacacacacacacacacacacacac 618
Db 403 ccacacacacacacacacacacacacacacacacacacacacacacacacacacac 462
Qy 619 ccacacacacacacacacacacacacacacacacacacacacacacacacacacac 678
Db 463 ccacacacacacacacacacacacacacacacacacacacacacacacacacacac 522
Qy 679 ctgctgtgacagacacacacacacacacacacacacacacacacacacacacac 738
Db 523 ctgctgtgacagacacacacacacacacacacacacacacacacacacacacac 582
Qy 739 ccttcctgacgtcctcacaagacacacacacacacacacacacacacacacacac 798
Db 583 ccttcctgacgtcctcacaagacacacacacacacacacacacacacacacacac 642
Qy 799 ggcacgtgacacacacacacacacacacacacacacacacacacacacacacac 858
Db 643 ggcacgtgacacacacacacacacacacacacacacacacacacacacacacac 702
Qy 859 ggcacgtgacacacacacacacacacacacacacacacacacacacacacacac 918
Db 703 ggcacgtgacacacacacacacacacacacacacacacacacacacacacacac 762
Qy 919 ctcatgtgacacacacacacacacacacacacacacacacacacacacacacac 978
Db 763 ctcatgtgacacacacacacacacacacacacacacacacacacacacacacac 822
Qy 979 ctgctgtgacacacacacacacacacacacacacacacacacacacacacacac 1038
Db 823 ctgctgtgacacacacacacacacacacacacacacacacacacacacacacac 882
Qy 1039 cacaacacacacacacacacacacacacacacacacacacacacacacacacacac 1098
Db 883 cacaacacacacacacacacacacacacacacacacacacacacacacacacacac 942
Qy 1099 aagccttgacacacacacacacacacacacacacacacacacacacacacacac 1158
Db 943 aagccttgacacacacacacacacacacacacacacacacacacacacacacac 1002
Qy 1159 ctgacacacacacacacacacacacacacacacacacacacacacacacacacacac 1218
Db 1003 ctgacacacacacacacacacacacacacacacacacacacacacacacacacacac 1062
Qy 1219 tagctgtgacacacacacacacacacacacacacacacacacacacacacacac 1278
Db 1063 tagctgtgacacacacacacacacacacacacacacacacacacacacacacac 1122
Qy 1279 gtaattcccaagatgaaagacacacacacacacacacacacacacacacacac 1338
Db 1279 gtaattcccaagatgaaagacacacacacacacacacacacacacacacacac
```

```
Db 1123 gtaattcccaagatgaaagacacacacacacacacacacacacacacacacac 1182
Qy 1339 tgcaaacacacacacacacacacacacacacacacacacacacacacacacacacacac 1398
Db 1183 tgcaaacacacacacacacacacacacacacacacacacacacacacacacacacacac 1242
Qy 1399 gacagacacacacacacacacacacacacacacacacacacacacacacacacacac 1458
Db 1243 gacagacacacacacacacacacacacacacacacacacacacacacacacacacac 1302
Qy 1459 ccttccttccttccttccttccttccttccttccttccttccttccttccttccttccttc 1518
Db 1303 ccttccttccttccttccttccttccttccttccttccttccttccttccttccttccttc 1362
Qy 1519 cacaagacacacacacacacacacacacacacacacacacacacacacacacacacacac 1578
Db 1363 cacaagacacacacacacacacacacacacacacacacacacacacacacacacacacac 1422
Qy 1579 acccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 1638
Db 1423 acccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 1482
Qy 1639 ctgactgtccttgaggaagacacacacacacacacacacacacacacacacac 1662
Db 1483 ctgactgtccttgaggaagacacacacacacacacacacacacacacacacac 1506
```

RESULT 7

US-09-687-527-407

Sequence 407, Application US/09687527

GENERAL INFORMATION:

APPLICANT: Yang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Zhou, Ping

APPLICANT: Asundi, Vinod

APPLICANT: Ren, Feiyan

APPLICANT: Zhang, Jie

APPLICANT: Zhao, Qing A.

APPLICANT: Xue, Aigong J.

APPLICANT: Yang, Tonghong

APPLICANT: Wehrman, Tom

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: Novel Nucleic Acids and

TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 795

CURRENT APPLICATION NUMBER: US/09/687,527

CURRENT FILING DATE: 2000-10-12

NUMBER OF SEQ ID NOS: 482

SOFTWARE: PL_Fl_genes Version 2.0

SEQ ID NO 407

LENGTH: 1537

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (184)...

US-09-687-527-407

Query Match 79.7%: Score 1324.8; DB 27; Length 1537;
Best Local Similarity 99.1%; Pred. No. 2.4e-285;
Matches 1332; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy 319 atcctaccatgacaggtgaaagacacacatgttgagaaatgacataagacaacct 378
Db 193 accctgacgtcctcacaagtgaagacagacatgtggaatgacataagacaacct 252
Qy 379 agccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 438
Db 253 agccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 312
Qy 439 agccttgacgtcctcacaagacacacacacacacacacacacacacacacacac 498
Db 439 agccttgacgtcctcacaagacacacacacacacacacacacacacacacacac
```


Db 1 gaagctggccagcgctgctacgacgctctgcgcgccttggttgccgctgcaagcacac 60
QY 1044 gaaagcgctcatctgcaaaagacagcgaggtgcaagacacagccgcaagacc 1103
Db 61 gaacagcgctcatctgcaaaagacagcgaggtgcaagacacagccgcaagacc 120
QY 1104 ttcttgacagtttgctgcaagcccgagctgctgcctgcaaaccttgctctctgac 1163
Db 121 ttcttgacagtttgctgcaagcccgagctgctgcctgcaaaccttgctctctgac 180
QY 1164 atctaagaagactgcaagaagcttgaaggttttagcctctgaagcttgtaact 1223
Db 181 atctaagaagactgcaagaagcttgaaggttttagcctctgaagcttgtaact 240
QY 1224 tgcacactccctaccccccagcttcggaataacagacacacacagctacactgta 1283
Db 241 tgcacactccctaccccccagcttcggaataacagacacacacagctacactgta 300
QY 1284 ccccaagtgatgaagaagacgttggtgttttttaaggttcctgaacttgta 1343
Db 301 ccccaagtgatgaagaagacgttggtgttttttaaggttcctgaacttgta 360
QY 1344 aacagacatgcaaggcgaggtgctgctggggggaatttctcttcttcaagaagacag 1403
Db 361 aacagacatgcaaggcgaggtgctgctggggggaatttctcttcttcaagaagacag 420
QY 1404 aacagacatgcaaggcgaggtgctgctggggggaatttctcttcttcaagaagacag 1463
Db 421 aacagacatgcaaggcgaggtgctgctggggggaatttctcttcttcaagaagacag 480
QY 1464 ctctccctccctccctcccgccccctctctcttcttcaatgcttctgcaactcaag 1523
Db 481 ctctccctccctccctcccgccccctctctcttcttcaatgcttctgcaactcaag 540
QY 1524 ggcctagctgctggggaagttgtaactgagttacagggtaactttaaagaagacct 1583
Db 541 ggcctagctgctggggaagttgtaactgagttacagggtaactttaaagaagacct 600
QY 1584 tggagctctctatacctctctctctccctccatccatccatccacccacttgctccgat 1643
Db 601 tggagctctctatacctctctctctccctccatccatccatccacccacttgctccgat 660
QY 1644 gctctggggaaggtgtaga 1662
Db 661 gctctggggaaggtgtaga 679

RESULT 12
US-09-943-143-34213
; Sequence 34213, Application US/09943143
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 783
; CURRENT APPLICATION NUMBER: US/09/943, 143
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/489, 036
; NUMBER OF SEQ ID NOS: 35324
; SOFTWARE: HY-patent.pl Version 3.1
; SEQ ID NO 34213
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-943-143-34213

Query Match 40.7%; Score 675.8; DB 35; Length 680;
Best Local Similarity 99.7%; Pred. No. 1.9e-140;
Matches 677; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 984 gaagctggccagcgctgctacgacgctctgcgcgccttggttgccgctgcaagcacac 1043

Db 1 gaagctggccagcgctgctacgacgctctgcgcgccttggttgccgctgcaagcacac 60
QY 1044 gaaagcgctcatctgcaaaagacagcgaggtgcaagacacagccgcaagacc 1103
Db 61 gaacagcgctcatctgcaaaagacagcgaggtgcaagacacagccgcaagacc 120
QY 1104 ttcttgacagtttgctgcaagcccgagctgctgcctgcaaaccttgctctctgac 1163
Db 121 ttcttgacagtttgctgcaagcccgagctgctgcctgcaaaccttgctctctgac 180
QY 1164 atctaagaagactgcaagaagcttgaaggttttagcctctgaagcttgtaact 1223
Db 181 atctaagaagactgcaagaagcttgaaggttttagcctctgaagcttgtaact 240
QY 1224 tgcacactccctaccccccagcttcggaataacagacacacacagctacactgta 1283
Db 241 tgcacactccctaccccccagcttcggaataacagacacacacagctacactgta 300
QY 1284 ccccaagtgatgaagaagacgttggtgttttttaaggttcctgaacttgta 1343
Db 301 ccccaagtgatgaagaagacgttggtgttttttaaggttcctgaacttgta 360
QY 1344 aacagacatgcaaggcgaggtgctgctggggggaatttctcttcttcaagaagacag 1403
Db 361 aacagacatgcaaggcgaggtgctgctggggggaatttctcttcttcaagaagacag 420
QY 1404 aacagacatgcaaggcgaggtgctgctggggggaatttctcttcttcaagaagacag 1463
Db 421 aacagacatgcaaggcgaggtgctgctggggggaatttctcttcttcaagaagacag 480
QY 1464 ctctccctccctccctcccgccccctctctcttcttcaatgcttctgcaactcaag 1523
Db 481 ctctccctccctccctcccgccccctctctcttcttcaatgcttctgcaactcaag 540
QY 1524 ggcctagctgctggggaagttgtaactgagttacagggtaactttaaagaagacct 1583
Db 541 ggcctagctgctggggaagttgtaactgagttacagggtaactttaaagaagacct 600
QY 1584 tggagctctctatacctctctctctccctccatccatccatccacccacttgctccgat 1643
Db 601 tggagctctctatacctctctctctccctccatccatccatccacccacttgctccgat 660
QY 1644 gctctggggaaggtgtaga 1662
Db 661 gctctggggaaggtgtaga 679

RESULT 13
US-09-457-877-66/C
; Sequence 66, Application US/09457877
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Contigs Obtained
; FILE REFERENCE: 20411-779
; CURRENT APPLICATION NUMBER: US/09/457, 877
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: HY-patent.pl Version 3.1
; SEQ ID NO 66
; LENGTH: 658
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)...(459)
; OTHER INFORMATION: similar to V32399|V32399 in the geneseq database. Run with
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51)...(459)
; OTHER INFORMATION: similar to gb|AF039843|AF039843 in the nt_19990223 databas

```

OTHER INFORMATION: e, Run with BLASTN 2.0a19MP-WashU, default parameters
FEATURE:
NAME/KEY: misc-feature
LOCATION: (3)...(365)
OTHER INFORMATION: similar to gb|AA499432|AA499432 in the dbEST, weekly, FASTA.
OTHER INFORMATION: 021899 database, Run with BLASTN 2.0a19MP-WashU, default par
FEATURE:
NAME/KEY: source
LOCATION: (235)...(659)
OTHER INFORMATION: Variant: 538088, Cluster: 582068, Library: adt.brain BRAIN G
OTHER INFORMATION: Lemn Fu
FEATURE:
NAME/KEY: source
LOCATION: (1)...(452)
OTHER INFORMATION: Variant: 779832, Cluster: 717962, Library: thymus THYMUS GF/
OTHER INFORMATION: MM
US-09-457-877-66

```

```

Query Match          39.6%; Score 658; DB 18; Length 658;
Best Local Similarity 100.0%; Pred. No. 1.8e-136;
Matches 658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

462 tttccaccacatctggaatctctcctcagcggcgcccgagctctgtgagcagcagcag 521
|||||
658 tctacacccacatctggaatctctcctcagcggcgcccgagctctgtgagcagcagcag 599
|||||
522 cacatctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 581
|||||
598 cactctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 539
|||||
582 ctacacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaac 641
|||||
538 ctacacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaac 479
|||||
642 gggcgccgagctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaac 701
|||||
478 gggcgccgagctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaac 419
|||||
702 gtgtaaatgcaagagtgatgcatcccccggagagctcttgaacgaagctcttgaacga 761
|||||
418 gtgtaaatgcaagagtgatgcatcccccggagagctcttgaacgaagctcttgaacga 359
|||||
762 ggaagtccttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 821
|||||
358 ggaagtccttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 299
|||||
822 gggcaatcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 881
|||||
298 gggcaatcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 239
|||||
882 ctccctgcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 941
|||||
238 ctccctgcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 179
|||||
942 gctcgccttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 1001
|||||
178 gctcgccttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 119
|||||
1002 ctacacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaac 1061
|||||
118 ctacacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaac 59
|||||
1062 agcagcagcggggaatgcaagaagcagcagcggggaagccttctgacagcttgg 1119
|||||
58 agcagcagcggggaatgcaagaagcagcagcggggaagccttctgacagcttgg 1
|||||

```

```

RESULT 14
US-09-952-981-66/c
Sequence 66, Application US/09952981
GENERAL INFORMATION:

```

```

APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Contigs Obtained
FILE OF INVENTION: From Various Libraries
FILE REFERENCE: 20411-779
CURRENT APPLICATION NUMBER: US/09/952,981
CURRENT FILING DATE: 2001-09-14
PRIORITY APPLICATION NUMBER: US/09/457,877
PRIORITY FILING DATE: 1999-12-08
NUMBER OF SEQ ID NOS: 128
SOFTWARE: Hy-patent.pl Version 3.1
SEQ ID NO 66
LENGTH: 658
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (51)...(459)
OTHER INFORMATION: similar to V323991V32399 in the geneseq database, Run
OTHER INFORMATION: BLASTN 2.0a19MP-WashU, default parameters
NAME/KEY: misc-feature
LOCATION: (51)...(459)
OTHER INFORMATION: similar to gb|AF039843|AF039843 in the nt_19990223
OTHER INFORMATION: databases
OTHER INFORMATION: e, Run with BLASTN 2.0a19MP-WashU, default parameters
NAME/KEY: misc-feature
LOCATION: (3)...(365)
OTHER INFORMATION: similar to gb|AA499432|AA499432 in the
OTHER INFORMATION: dbEST, weekly, FASTA.
OTHER INFORMATION: 021899 database, Run with BLASTN 2.0a19MP-WashU, default par
OTHER INFORMATION: ameters
NAME/KEY: source
LOCATION: (235)...(659)
OTHER INFORMATION: Variant: 538088, Cluster: 582068, Library: adt.brain
OTHER INFORMATION: BRAIN G
OTHER INFORMATION: Lemn Fu
NAME/KEY: source
LOCATION: (1)...(452)
OTHER INFORMATION: Variant: 779832, Cluster: 717962, Library: thymus
OTHER INFORMATION: THYMUS GF/
OTHER INFORMATION: MM
US-09-952-981-66

```

```

Query Match          39.6%; Score 658; DB 35; Length 658;
Best Local Similarity 100.0%; Pred. No. 1.8e-136;
Matches 658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

462 tttccaccacatctggaatctctcctcagcggcgcccgagctctgtgagcagcagcag 521
|||||
658 tctacacccacatctggaatctctcctcagcggcgcccgagctctgtgagcagcagcag 599
|||||
522 cacatctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 581
|||||
598 cactctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 539
|||||
582 ctacacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaac 641
|||||
538 ctacacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaac 479
|||||
642 gggcgccgagctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaac 701
|||||
478 gggcgccgagctcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaac 419
|||||
702 gtgtaaatgcaagagtgatgcatcccccggagagctcttgaacgaagctcttgaacga 761
|||||
418 gtgtaaatgcaagagtgatgcatcccccggagagctcttgaacgaagctcttgaacga 359
|||||
762 ggaagtccttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 821
|||||
358 ggaagtccttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 299
|||||
822 gggcaatcttgaacgaagctcttgaacgaagctcttgaacgaagctcttgaacgaag 881
|||||

```

```
Db 298 GGGGATCTTTACACATGCACGATGAGGACATGAGGCTCTCTGTCACACCCCTG 239
OY 882 CTCCTGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 941
Db 238 CTCCTGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 179
OY 942 GCTGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 1001
Db 178 GCTGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 119
OY 1002 CTCGACCTGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 1061
Db 118 CTCGACCTGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 59
OY 1062 AGCAGCAGCGGGGATGCCAAGACGACGACGACGACGACGACGACGACGACGAC 1119
Db 58 AGCAGCAGCGGGGATGCCAAGACGACGACGACGACGACGACGACGACGACGAC 1
```

RESULT 15

```
US-09-489-036-15011/c
; Sequence 15011, Application US/09489036
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 783
; CURRENT APPLICATION NUMBER: US/09/489,036
; NUMBER OF SEQ ID NOS: 35324
; SOFTWARE: Hy-Patent-PI Version 3.1
; SEQ ID NO 15011
; LENGTH: 635
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-036-15011
```

Query Match

```
Best Local Similarity 35.1%; Score 583.6; DB 18; Length 635;
Matches 620; Conservative 0; Mismatches 4; Indels 4; Gaps 3;
```

```
OY 53 CTTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 111
Db 626 CTTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 567
OY 112 TCAAGATTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 170
Db 566 TCAAGATTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 507
OY 171 AGAAGCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 230
Db 506 AGAAGCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 447
OY 231 GAGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 290
Db 446 GAGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 387
OY 291 CCACAGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 350
Db 386 CCACAGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 327
OY 351 TGTGGAATATGACTACATGACACACCTTACCTTACCTTACCTTACCTTACCTTACCTT 410
Db 326 TGTGGAATATGACTACATGACACACCTTACCTTACCTTACCTTACCTTACCTTACCTT 267
OY 411 CCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 470
Db 266 CCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 207
OY 471 CCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 530
```

```
Db 206 CCATTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 149
OY 531 TGACCAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 590
Db 148 TGACCAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 89
OY 591 GGTCTGCGATCAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 650
Db 88 GGTCTGCGATCAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 29
OY 651 GGTCTGCGATCAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 678
Db 28 GGTCTGCGATCAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1
```

Search completed: May 11, 2002, 06:19:13
Job time: 14498 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2002, 21:32:29 ; Search time 1500.46 Seconds

(without alignments)
14950.042 Million cell updates/sec

Title: US-10-082-902-1

Perfect score: 1662

Sequence: 1 agtagagcagctgagctgac.....tgctctggggaagctgtaga 1662

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_hic:*
13: em_gss_hum:*
14: em_gss_hiv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	498.2	30.6	679	9	BB648007
2	475.4	28.6	666	9	BB630302
3	383.6	23.1	448	10	BE709917
4	373	22.4	471	10	BE721232
5	308.6	18.6	456	9	AA499432
6	308	18.5	726	10	BF579947
7	279.8	16.8	377	9	AF643238
8	233.2	14.0	756	12	CNS023FM
9	222.6	13.4	557	9	AM128750
10	202.2	12.2	2486	11	AK014578
11	201.8	12.1	401	9	A1545888
12	201.8	12.1	657	12	AZ351672
13	195	11.7	1160	12	CNS03FRP
14	188.6	11.3	1023	10	BF580057
15	186.8	11.2	802	10	BI914916
16	186	11.2	562	10	BJ032161
17	181	10.9	584	9	AW959446

18	180.2	10.8	527	10	BE270351	BE270351 601186509
19	180.2	10.8	666	10	BE296059	BE296059 601177045
20	180.2	10.8	1719	11	BC004205	BC004205 Homo sapi
21	179.4	10.8	439	9	A1613504	A1613504 v156910.y
22	178.4	10.7	522	9	AW463200	AW463200 BP230012A
23	178.2	10.7	517	9	AW266965	AW266965 0266-5 Le
24	178.2	10.7	522	10	BF040166	BF040166 BP250022A
25	178.2	10.7	522	10	BF04187	BF04187 BP250012B
26	178.2	10.7	705	9	AV650436	AV650436 AV650436
27	177	10.6	531	10	BE295152	BE295152 601174909
28	172.4	10.4	720	9	A1652823	A1652823 AL652823
29	172.4	10.4	720	10	BC116037	BC116037 602318033
30	172.2	10.4	823	10	BF122111	BF122111 601755970
31	169.8	10.2	609	9	AA31912	AA31912 zw88a08.r
32	167.6	10.1	976	12	CNS04OJO	AL300093 Tetradon
33	164	9.9	710	10	BE294735	BE294735 601173968
34	160.4	9.7	688	10	BJ060732	BJ060732 BJ060732
35	150.6	9.1	552	9	AW955843	AW955843 EST367913
36	146.4	8.8	493	10	BF090636	BF090636 OV2-NT004
37	145.4	8.7	958	9	A1534238	AL534238 AL534238
38	140.6	8.5	542	9	BC030374	BC030374 128423 MA
39	137.8	8.3	388	10	BE934390	BE934390 ILS-NT027
40	137.6	8.3	732	10	BF152022	BF152022 602916061
41	135.4	8.1	368	9	AA367247	AA367247 EST78383
42	135.4	8.1	421	10	R69044	R69044 Y143109.r1
43	135.2	8.1	628	10	BG482909	BG482909 602502916
44	131.6	7.9	609	10	BE92466	BE92466 601434360
45	129.4	7.8	402	10	BF933262	BF933262 ILS-NT027

ALIGNMENTS

RESULT 1
BB648007
LOCUS
DEFINITION
BB648007 RIKEN full-length enriched, 16 days embryo head Mus
musculus cDNA clone C130036A21.5, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus

REFERENCE
AUTHORS
1 (bases 1 to 679)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

RIKEN House ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10): 1617-1630 (2000)
Wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LINT; contact the
IMAGE Consortium (info@image.lnlgov) for further information.
G81:532011

AI643238	AI643238	377 bp	MRNA	linear	EST 29-APR-1999
LOCUS					
DEFINITION	v19310.y1 Stratigene mouse heart (4937316) Mus musculus cDNA clone IMAGE:919795 5' similar to TR:043609 O43609 NOVEL ANTAGONIST OF FGF3 SIGNALLING ;, mRNA sequence.				
ACCESSION	AI643238				
VERSION	AI643238.1	GI:4721713			
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus.				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus. 1 (bases 1 to 377)				
REFERENCE	Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepic,M., Theisling,B., Allen,M., Bowers,Y., Person ,B., Swaller,T., Gibbons,M., Page,D., Harey,N., Schurk,R., Rlter ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterson,R. and Wilson,R.				
TITLE	The MASHU-NCI Mouse EST Project 1999				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Marra M/MASHU-NCI Mouse EST Project 1999				

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG15532011
This read is a RESSEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
Seq primer: -40RP from Glibco
High quality sequence stop: 376.

FEATURES
SOURCE

```

/organism="Mus musculus"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:919795"
/clone_id="Stratagene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/leh_host="SOLR (kanamycin resistant)"
/node_organ="heart; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI; Cloned unidirectionally. Primer
Oligo dT. 93 pooled NIH/Swiss 13 day embryo hearts.
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGCACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTTT 3'"
73 a 119 c 105 g 79 t 1 others

```

Query Match	16.8%	Score 279.8	DB 9	Length 377
Best local similarity	87.9%	Pred. 4.6e-49		
Matches 327	Conservative 0	Mismatches 43	Indels 2	Gaps 2
QY	755	gaaacaggagatgcctgtgtctcgaagccagactctgtcaacatctgacgtgacatgatt	814	
Db	1	GCAMCCAGAGATGCTGTGTCTGCTGCGCTCAGACGCTGGTCAACTGTGGCATATCAGATGTC	60	
QY	815	tggtgcaggagatctctacacactgcagaaatgaagacgaatgaaggctcctgcgtgacc	874	
Db	61	TGCTGCAAGGTATCTTCTATCATCTGACACTGATGAGATGATGAGGGCTCTTGCGCGACC	120	
QY	875	aacctctgctctgcctccgctcccaactctctggccgcctgtgtctctcatgggtgtctac	934	
Db	121	ACCCCTGCTCTGTTCCCGGCTCCAACTGCTGTGCGCCCGCTGTCTTCAATGAGGCGCCCTCT	180	
QY	935	cctgtgtgtgcctctgcgtctctgctactactgtcctctccacaggtctgctgaaagcttgccc	994	

[illegible]

RESULT	8
CNS023FM/C	
LOCUS	
DEFINITION	CNS023FM 756 bp DNA linear GSS 12-MAY-2000 Tetradon nigroviridis genome survey sequence T7 end of clone 22Bb4 of library G from Tetradon nigroviridis, genomic survey sequence.
ACCESSION	AL179419
VERSION	AL179419.1 GI:7817476
KEYWORDS	GSS; genome survey sequence, Tetradon nigroviridis,
SOURCE	Tetradon nigroviridis
ORGANISM	

REFERENCE
AUTHORS
Roest-Crolius, H., Jallón, O., Dasilva, C., Fizmaes, C., Fisher, C.,
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Telostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 756)

TITLE
Journal
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish *Tetraodon nigroviridis*
unpublished
2. (pages 1 to 756)
REFERENCE

<p> AUTHORS Roest-Collins, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Filames, C., Mincker, P., Brotlier, P., Quetier, F., Saurin, W. and Weissenbach, J. </p>	<p> TITLE Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence </p>
<p> JOURNAL Unpublished </p>	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
3	(bases 1 to 756)	Genoscope.	Direct Submission	Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigriviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon_location/dna/altifers

	source	1..756	/organism="Tetracodon nigroviridis"
		/db_xref="taxon:9883"	
		/clone="232B14"	
		/clone_1lb="G"	
		/note="Genoscope sequence ID : C0AG32DA07LPI-end : T7"	
BASP COUNT		133 a	210 c 265 g 142 t 6 others
ORIGIN			

Query Match	14.0%	Score 233.2	DB 12	Length 756
Best Local Similarity	73.9%	Pred. No. 4,3e-39		
Matches 308	Conservative 1	Mismatches 105	Indels 3	Gaps 1

QY	654	ccccaccgagctggaacaacattctctctgtcgagagcctgttgggaagtgtaatatgcaa	713
Db	527	cttccccggccggaagaacacactcttctctcttcggaacaactctcggaagccgacgac	468
QY	714	gagagtgtatcccccggagcttgcctctctctgtgtgttcacaacaggatgtcctgtg	773

Db 467 GGAAGTACGCTCCCGGAGACCCCTCCCTGCTGCTGCTGCAACCAAGAGAGTGCCTGTG 408
 Qy 774 ctcagcccaagctctgtcaactatgacagctgctgtgttggctgagggcatcttctca 833
 Db 407 CTCGGCCCAAGAGCTGTGGAGGAGGACCACTGCTGTGCTGTCAAGAGCATCTTCA 348
 Qy 834 ccaactgac---gaatgaagacgagatgaggtctctgctgctgacacccctgtctctgc 890
 Db 347 CCACTGGCAG 288
 Qy 891 ccagctcaactgctgcccgcctgtgtctctcaatgagtgagctgctgctgctgctgctg 950
 Db 287 CCAAGCCCAAGCTGTGGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 228
 Qy 951 cctgctctgtactgctgctgacagcagctgctgctgctgctgctgctgctgctgctg 1010
 Db 227 TCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 168
 Qy 1011 tctgcccgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1067
 Db 167 CGTCAGAGAGCGCGCTGCTGCTGCAAGACTGTGCAAGAGCGGAGTCAAGTCCCTGC 111

RESULT 9
 LOCUS AM128750 557 bp mRNA linear EST 07-JUN-2001
 DEFINITION fe37a01.y1 zebrafish Washu MPIMG EST Danio rerio cDNA clone
 IMAGE:3741000 5' similar to TR:043597 043597 SPROUTY 2. ; mRNA
 sequence.
 ACCESSION AM128750
 VERSION AM128750.1 GI:6116654
 KEYWORDS EST.
 SOURCE zebrafish.
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
 ; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 557)
 AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Merra,M., Eddy
 S., Hillier,L., Kuehba,T., Martin,J., Beck,C., Wylie,T., Underwood
 K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Peterson,B.,
 Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler,E.,
 Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
 and Wilson,R.
 TITLE Washu Zebrafish EST Project 1998
 JOURNAL Unpublished (1998)
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: Zbratfish@watson.wustl.edu
 CDNA Library Preparation: Matthew Clark. CDNA Library Arrayed by:
 Matthew Clark. DNA Sequencing by: Washington University Genome
 Sequencing Center Clone Distribution: Genome Systems, St. Louis,
 Missouri (web address: www.genomesystems.com) (email contact:
 info@genomesystems.com) and Research Genetics, Huntsville, Alabama
 (web address: www.resgen.com) (email contact: info@resgen.com) and
 RessourcenzentrumPirmasendebank, Berlin, Germany (web address:
 www.rzp.de)
 Seq primer: T3 ET from Amersham
 High quality sequence stop: 494.

FEATURES
 source
 Location/Qualifiers
 1..557

/organism="Danio rerio"
 /db_xref="taxon:7955"
 /clone="IMAGE:3741000"
 /clone.lib="Zebrafish Washu MPIMG EST"
 /sex="mixed"
 /tissue.type="26 somite embryos, adult livers, shield
 stage embryos"
 /lab.host="XLI-blue MRF"
 /note="Vector: pSPORT1, Site_1: NotI; Site_2: SalI; 1st

BASE COUNT 133 a 147 c 147 g 130 t
 ORIGIN

Query Match 13.4%. Score 222.6; DB 9; Length 557;
 Best Local Similarity 72.5%; Pred. No. 6.7e-37;
 Matches 288; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 647 cagcagctccaccagcagctgagcaagcactctgtgtgctgagcctgtggaagtta 706
 Db 5 CACGCTTGCTGTGAG 64
 Qy 707 aatgcaagagatgtgcatcccccggagctgtgcttctgtgctgtgcaacagagt 766
 Db 65 GATGACCGCAAGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 124
 Qy 767 gctctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 826
 Db 125 GTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 184
 Qy 827 tctctac 886
 Db 185 TTTTCTACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244
 Qy 887 gctccgcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 946
 Db 245 GTTCGACACTGCAAGCTGTGCGCAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 304
 Qy 947 cctgctgctgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1006
 Db 305 CCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 364
 Qy 1007 accgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1043
 Db 365 ACGGCTGACG 401

RESULT 10
 LOCUS AK014578

DEFINITION AK014578 2486 bp mRNA linear HTC 19-JAN-2002
 Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched
 library, clone:4632415G13:sprouty homolog 1 (Drosophila), full
 insert sequence.

ACCESSION AK014578
 VERSION AK014578.1 GI:12852518
 KEYWORDS HTC, CAP trapper.
 SOURCE Mus musculus (Strain:C57BL/6J) 0 day neonate skin cDNA to mRNA,
 clone.lib:RIKEN full-length enriched mouse cDNA library
 clone:4632415G13.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE Carninci,P. and Hayashizaki,Y.
 1 (sites)
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)


```
Db 287 CAACGAGACAGTCCCTGGTCTAGTCTGAAAGGCATCCGACCCAGCCAGACAC 346
QY 617 tcaactgcagccgctggaactcaaggcccgcggtcccaaccgagcttgacaagcaact 676
Db 347 T-----GATTGGATGACTTGAAGGGTCTCTTGAAGAAGAGACCTGACACACACA 397
QY 677 tcttgctgtagaagacctggaagttaaatgcaagagtgatgacatcccccggagct 736
Db 398 AGTTCAATTGTGAAACAGTGTGGAGTGCAGATGTGGAGAAATGCACTGCTCCAGGACCC 457
QY 737 tgcctccctgcctgggtctgcaaacagagtgctgctcagcccaagactctggtcaact 796
Db 458 TACCATCTCTGTGGCTGTAAACGGCACTGCTTGTCTGCTGAGAGCATGTGTGAAT 517
QY 797 atggcagctgcatgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 856
Db 518 ATGGAACCTGCATGTGCTTAGTCAAGGGCATCTTCTACACCTGCTCCAAATGAGAGGAA 577
QY 857 agggctcctgcctgaccccccctgctcctgctcctgctcctgctcctgctgctgct 916
Db 578 GGGATTCTTATTCAGTAATCTTCTCTCTCTTACAAATCACACTGCTGCTAGATACC 637
QY 917 ccttcagggtgctcctcctcctgctgctgctgctgctgctgctgctgctgctgct 976
Db 638 TGTGTATGGAGCCATGCTTATTTTACCTTACTGCTGTATCTCTCTCTGCTAAG 697
QY 977 gctgcgtgaagctgagcccgctgctgctgctgctgctgctgctgctgctgctgctg 1036
Db 698 GATGCTGAAGCTGTGCAAGAGTGTCTATGACTGATCATCGCCAGGGTGCAGATGA 757
QY 1037 agcacagaaacagcgatcctgcaagcag 1066
Db 758 AGAACTCCAAACACTTGTCTATTTGTAAGCTG 787
```

Search completed: May 11, 2002, 02:42:24
Job time: 18595 sec

[illegible]

Db 301 ctcagacacccatcacatctccatctgacaggtgaaagcagccatgtgaaagat 360
Qy 361 gactaatagaacacccatagcttgccctgacacccagccaaagcgacccggcg 420
Db 361 gactaatagaacacccatagcttgccctgacacccagccaaagcgacccggcg 420
Qy 421 gccacgaatctgcccagccagcccgccgctgacacgaatgtacacacatgttc 480
Db 421 gccacgaatctgcccagccagcccgccgctgacacgaatgtacacacatgttc 480
Qy 481 tccctgaagcgagccacagctctgtgacagcagcagcagcactctctgacacag 540
Db 481 tccctgaagcgagccacagctctgtgacagcagcagcagcactctctgacacag 540
Qy 541 ctcctgacacatgac 600
Db 541 ctcctgacacacatgac 600
Qy 601 atccagcccaaggtggtccacgtccagccgctggaacctcaagggccgggtccacac 660
Db 601 atccagcccaaggtggtccacgtccagccgctggaacctcaagggccgggtccacac 660
Qy 661 gaagctggaagacactctctgtgacagccgctgtggaaggttaagcagagtggt 720
Db 661 gaagctggaagacactctctgtgacagccgctgtggaaggttaagcagagtggt 720
Qy 721 gaatcccccagagctgctcctcctgtggtctgtcaacagagagtgctgtgacgc 780
Db 721 gaatcccccagagctgctcctcctgtggtctgtcaacagagagtgctgtgacgc 780
Qy 781 caaacctgtgtaactatgacagcagctgtgttctgtgacagcactctctacacac 840
Db 781 caaacctgtgtaactatgacagcagctgtgttctgtgacagcactctctacacac 840
Qy 841 acgaatgagagcagatgagaggtctctgtgctgtgacacacacacacacacacac 900
Db 841 acgaatgagagcagatgagaggtctctgtgctgtgacacacacacacacacacac 900
Qy 901 tgcctgacccgctgctcctcctcctcctcctcctcctcctcctcctcctcctc 960
Db 901 tgcctgacccgctgctcctcctcctcctcctcctcctcctcctcctcctcctc 960
Qy 961 taactgctgac 1020
Db 961 taactgctgac 1020
Qy 1021 cctgtgtgctgacagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
Db 1021 cctgtgtgctgacagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
Qy 1081 aagaacagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
Db 1081 aagaacagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
Qy 1141 tggaaacacgtgtctctcctcctcctcctcctcctcctcctcctcctcctcctc 1200
Db 1141 tggaaacacgtgtctctcctcctcctcctcctcctcctcctcctcctcctcctc 1200
Qy 1201 tccctgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260
Db 1201 tccctgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260
Qy 1261 accac 1320
Db 1261 accac 1320
Qy 1321 cagaggtctcctgaaac 1380
Db 1321 cagaggtctcctgaaac 1380
Qy 1381 aattttctcttctcagaaagacagaaacacacacacacacacacacacacacac 1440
Db 1381 aattttctcttctcagaaagacagaaacacacacacacacacacacacacacac 1440

Db 1381 aattttctcttctcagaaagacagaaacacacacacacacacacacacacacac 1440
Qy 1441 cttgaatgctctccacagccctcctcctcctcctcctcctcctcctcctcctcct 1500
Db 1441 cttgaatgctctccacagccctcctcctcctcctcctcctcctcctcctcctcct 1500
Qy 1501 tccatgtctcttgac 1560
Db 1501 tccatgtctcttgac 1560
Qy 1561 agggatcctttaaagaagacacacacacacacacacacacacacacacacacacac 1620
Db 1561 agggatcctttaaagaagacacacacacacacacacacacacacacacacacacac 1620
Qy 1621 ctcac 1680
Db 1621 ctcac 1680
RESULT 2
PCT-US02-08781-2
; Sequence 2, Application PC/TUS0208781
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Tom Y.
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Liu, Chenghua
; APPLICANT: Weinman, Tom
; APPLICANT: Ren, Feiyan
; APPLICANT: Dirmann, Radoje, T.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-006-061 CIP
; CURRENT APPLICATION NUMBER: PCT/US02/08781
; PRIOR APPLICATION NUMBER: 09/814,354
; PRIOR FILING DATE: 2001-3-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 2061
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/REV: CDS
; LOCATION: (245)..(1144)
; PCT-US02-08781-2
Query Match 98.7%; Score 1640; DB 1; Length 2061;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
Qy 3 gtagcagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 62
Db 37 gtagcagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 96
Qy 63 tccctcagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 122
Db 97 tccctcagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 156
Qy 123 acagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 182
Db 157 acagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 216
Qy 183 ctcctgac 242
Db 217 ctcctgac 276
Qy 243 gactcccaactatgac 302
Db 277 gactcccaactatgac 336

Query Match 88.6%; Score 1472.2; DB 1; Length 5686;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1503; Conservative 0; Mismatches 13; Indels 3; Gaps 2;

QY 147 cccctgagctctctcaagagcccttagaagcctgttctctcgtacagtcagacccacag 206
 Db 956 ccccaaccaccccccagagcccttagaagcctgttctctcgtacagtcagacccacag 1015

QY 207 ccccatgagagcccccagatcccaagagccccccttgatccccaactcagtcagtcaca 266
 Db 1016 ccccaagagagcccccagatcccaagagccccccttgatccccaactcagtcagtcaca 1075

QY 267 gccctctcttgacagagcggatgtcccaagcggctcccaagcacaactcacaactcacc 326
 Db 1076 gccctctcttgacagagcggatgtcccaagcggctcccaagcacaactcacaactcacc 1135

QY 327 cattgagccaggtgagagacagacagatgtgagatgactataagacaacactagaccctgac 386
 Db 1136 cattgagccaggtgagagacagacagatgtgagatgactataagacaacactagaccctgac 1195

QY 387 cctgacacacagcccaagcagagccgggagggcccaagcctgagccccaagccagccagc 446
 Db 1196 cctgacacacagcccaagcagagccgggagggcccaagcctgagccccaagccagccagc 1255

QY 447 ccgctgtgacagagatgtacacacacatgtgactctcttagagggccgcccagagcctgtg 505
 Db 1256 ccgctgtgacagagatgtacacacacatgtgactctcttagagggccgcccagagcctgtg 1315

QY 506 tggagcagacagag--cagacaactccttgacacaaagagcctcttagaacacaaatggacacac 563
 Db 1316 tggagcagcagcagacagacacacactctcttgacacaaagagcctctcttagaacacaaatggacacac 1375

QY 564 acccgctgacagcagagcctcacaagagagcctgtgacatccacaacccaaggtgtgtccactgt 623
 Db 1376 acccgctgacagcagagcctcacaagagagcctgtgacatccacaacccaaggtgtgtccactgt 1435

QY 624 ccagccagcctgagaccccaagagggccgggtccccaacccagagcctgagacagacactctgtcgt 683
 Db 1436 ccagccagcctgagaccccaagagggccgggtccccaacccagagcctgagacacactctctgtcgt 1495

QY 684 gtgagagagcctgtgagaaagtgtanaatgaaagagtgatgtgacatcccccagagctgtccttc 743
 Db 1496 gtgagagagcctgtgagaaagtgtanaatgaaagagtgatgtgacatcccccagagctgtccttc 1555

QY 744 ctgctgagcctgtcacaacccagagagtgccgtgtcagcccaagacccctgttcaactatggagac 803
 Db 1556 ctgctgagcctgtcacaacccagagagtgccgtgtcagcccaagacccctgttcaactatggagac 1615

QY 804 gtgacatgtgttctgagcagagacatctctacacactgacagaaatgaaagagagagcctc 863
 Db 1616 gtgacatgtgttctgagcagagacatctctacacactgacagaaatgaaagagagagcctc 1675

QY 864 ctgagcgtgagcacaacccctctctctctcctccgctcccaactgtctgagccgctgttcttcaat 923
 Db 1676 ctgagcgtgagcacaacccctctctctctcctccgctcccaactgtctgagccgctgttcttcaat 1735

QY 924 ggggtgctctctcctgagcgtgagcctgtcgtctcgtctacactgctgacacccagagctgct 983
 Db 1736 ggggtgctctctcctgagcgtgagcctgtcgtctcgtctacactgctgacacccagagctgct 1795

QY 984 gaaagcgtgagcagagctgtgctacagacgctctgagcagcctgtgtctgagctgacagacac 1043
 Db 1796 gaaagcgtgagcagagctgtgctacagacgctctgagcagcctgtgtctgagctgacagacac 1855

QY 1044 gaaacagctgcatactgtcacaagcagacagcgggagatgcacaagacagagacagccggaagac 1103
 Db 1856 gaaacagctgcatactgtcacaagcagacagcgggagatgcacaagacagagacagccggaagac 1915

QY 1104 ttcttgacagctgtgtctgaaagcccaagtgctctgctgtgaaactgtgtctctctctgac 1153
 Db 1916 ttcttgacagctgtgtctgaaagcccaagtgctctgctgtgaaactgtgtctctctctgac 1975

QY 1164 atctaaagagctgacagcaaggtcagaaggttttagcctctcgtgagcctgtgactagtc 1223
 Db 1976 atctaaagagctgacagcaaggtcagaaggttttagcctctcgtgagcctgtgactagtc 2035

QY 1224 tgcacatccctcaccacccagcttcgagaaatatacagagagacacacacacagctacgtgtatt 1283
 Db 2036 tgcacatccctcaccacccagcttcgagaaatatacagagagacacacacacagctacgtgtatt 2095

QY 1284 ccccaagatgatatgaagaagacacttgggctttttctcagaggtcctgaaactgtgtca 1343
 Db 2096 ccccaagatgatatgaagaagacacttgggctttttctcagaggtcctgaaactgtgtca 2155

QY 1344 aacagacatgacagagggcagaggtgtggtttggggggaataattctctcttttcagaagacag 1403
 Db 2156 aacagacatgacagagggcagaggtgtggtttggggggaataattctctcttttcagaagacag 2215

QY 1404 aacagagatgtgagacacatcccgaaactgcagcgtctgatatgctctccagccctc 1453
 Db 2216 aacagagatgtgagacacatcccgaaactgcagcgtctgatatgctctccagccctc 2275

QY 1464 ctctccctccctccctccctccgccccccctctctctctcttccatgtctcttgagactcaag 1523
 Db 2276 ctctccctccctccctccctccgccccccctctctctctcttccatgtctcttgagactcaag 2335

QY 1524 gaggctagctgctgagggagaggtgtactagatgactacagaggtactttaagaagacacct 1583
 Db 2336 gaggctagctgctgagggagaggtgtactagatgactacagaggtactttaagaagacacct 2395

QY 1584 tggagctctctatacctctctctctctcccaactcactcaccacacacacttgtccctgtat 1643
 Db 2396 tggagctctctatacctctctctctctcccaactcactcaccacacacacttgtccctgtat 2455

QY 1644 gtcttgggggagaggtgtaga 1662
 Db 2456 gtcttgggggagaggtgtaga 2474

RESULT 4
 US-10-116-802-373
 ; Sequence 373, Application US/10116802
 ; GENERAL INFORMATION:
 ; APPLICANT: Amy Lasek
 ; TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
 ; FILE REFERENCE: PA-0045 US
 ; CURRENT APPLICATION NUMBER: US/10/116,802
 ; PRIOR FILING DATE: 2002-04-04
 ; PRIOR APPLICATION NUMBER: 60/281,593
 ; PRIOR FILING DATE: 2001-04-04
 ; NUMBER OF SEQ ID NOS: 519
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 373
 ; LENGTH: 3475
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No: 241742.1
 ; NAME/KEY: unsure
 ; LOCATION: 1968
 ; OTHER INFORMATION: a, t, c, g, or other
 US-10-116-802-373

Query Match 13.9%; Score 230.2; DB 6; Length 3475;
 Best Local Similarity 98.7%; Pred. No. 8,1e-43;
 Matches 232; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1428 gaaactgagctgtctgtatgacttcccaagccctctctctcctccctccctccccc 1487
 Db 1 gaaactgagctgtctgtatgacttcccaagccctctctctcctccctccctccccc 60

QY 1488 ccccttcccttctcactgtcttcttgacactcagagagctagcgtcctgggaggaattg 1547

Best Local Similarity 66.7%; Pred. No. 1.4e-31;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

```

Qy 661 gagctggaagaacgactcttgcctgctgagagccttgggaagtgaatgaagaagtgt 720
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1223 GATTGGGCGCTGCACGCCCTACAGGTGTGAGAGCTGTGGCAAGTGAAGTGAAGAGTGC 1164
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 721 gcatcccccgcgaagcttgccttctctgctgctgcaacagagagtgctgctgcaac 780
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1163 ACCTACCCAAAGGCGCTGTGCATGACACTGATCTGCGAAGCAAGCAAGTGCCTTGTGCGCC 1104
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 781 cagactctgtcaactatgacagctgcatgctgttgggtgagcaggaattcttaccatgc 840
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1103 CAGAACGATGATGTAGTGGACTTGTGTATGCTGTGTAAGGCTCTTCTTATACACTGT 1044
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 841 acgaatgagagcagatgagggctctctgctgacacccctgctctctgctccgctccac 900
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1043 TCTAAT---GATGATGAGGACAACTGTGCTGACAACTGATGTTCTTGAAGCAAGTCTCAC 987
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 901 tgcctgcccgcctgctctctcatgagtgctctctccgctgagtgctgctgctgctctgc 960
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 986 TGTGTACAGAGAGTACGACATGGGTGATGATGCTCTTCTTCTTCTTCTTCTTCTTCTTCT 927
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 961 taactgctgctgacagcagctgctgctgctgctgctgctgctgctgctgctgctgctgc 1020
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 TACCTTCACAGCAAGGCTTCCCTTAATTGTGCTGCAAGGGGTCTTATGACCGGGTTACACG 867
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1021 cctgtgtgcccgcctgcaagcacagcagctgcatctgcaagcagcga 1069
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 CCTGTGCGCGCTGTAAATACTAAACACAGTTTGTCTGCAAGTCCCA 818
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 11

US-10-105-299-13407/c
; Sequence 13407, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105, 299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 13407
; LENGTH: 4802
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-13407

Query Match 10.8%; Score 180.2; DB 6; Length 4802;
Best Local Similarity 66.7%; Pred. No. 1.4e-31;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

```

Qy 661 gagctggaagaacgactcttgcctgctgagagccttgggaagtgaatgaagaagtgt 720
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1223 GATTGGGCGCTGCACGCCCTACAGGTGTGAGAGCTGTGGCAAGTGAAGTGAAGAGTGC 1164
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 721 gcatcccccgcgaagcttgccttctctgctgctgcaacagagagtgctgctgcaac 780
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1163 ACCTACCCAAAGGCGCTGTGCATGACACTGATCTGCGAAGCAAGCAAGTGCCTTGTGCGCC 1104
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 781 cagactctgtcaactatgacagctgcatgctgttgggtgagcaggaattcttaccatgc 840
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1103 CAGAACGATGATGTAGTGGACTTGTGTATGCTGTGTAAGGCTCTTCTTATACACTGT 1044
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 841 acgaatgagagcagatgagggctctctgctgacacccctgctctctgctccgctccac 900
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1043 TCTAAT---GATGATGAGGACAACTGTGCTGACAACTGATGTTCTTGAAGCAAGTCTCAC 987
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 901 tgcctgcccgcctgctctctcatgagtgctctctccgctgagtgctgctgctgctgctgc 960
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 986 TGTGTACAGAGATGTCAGCACATGGGTGTATGCTCTTCTTCTTCTTCTTCTTCTTCTTCT 927
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 961 taactgctgctgacagcagctgctgctgctgctgctgctgctgctgctgctgctgctgc 1020
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 TACCTTCACAGCAAGGCTTCCCTTAATTGTGCTGCAAGGGGTCTTATGACCGGGTTACACG 867
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1021 cctgtgtgcccgcctgcaagcacagcagctgcatctgcaagcagcga 1069
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 CCTGTGCGCGCTGTAAATACTAAACACAGTTTGTCTGCAAGTCCCA 818
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 12

US-10-105-299-8309
; Sequence 8309, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105, 299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 8309
; LENGTH: 5325
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-8309

Query Match 10.8%; Score 180.2; DB 6; Length 5325;
Best Local Similarity 66.7%; Pred. No. 1.5e-31;
Matches 273; Conservative 0; Mismatches 133; Indels 3; Gaps 1;

```

Qy 661 gagctggaagaacgactcttgcctgctgagagccttgggaagtgaatgaagaagtgt 720
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4103 gatttggtctgcaagcagctgcaagagtgctgagagctgctgcaagtgcaaatgaaggagtg 4162
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 721 gcatcccccgcgaagcttgccttctctgctgctgctgctgctgctgctgctgctgctgctgc 780
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4163 acccaaccaagagcctctgcatcatgagctgcatctgcaagcagtgcttctgctgctgc 4222
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 781 cagactctgtcaactatgacagctgcatgctgttgggtgagcaggaattcttaccatgc 840
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4223 cagaacgtgcatgactatgagctgtgctgctgctgctgctgctgctgctgctgctgctgctgc 4282
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 841 acgaatgagagcagatgagggctctctgctgctgctgctgctgctgctgctgctgctgctgctgc 900
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4283 tctaatt---gagtatgagagcaactgtgctgcaaacacatgtcttctgaagcagcttccac 4339
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 901 tgcctgcccgcctgctctctcatgagtgctctctccgctgagtgctgctgctgctgctgc 960
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4340 tgtgtacagagatgctgcaagcagatggtgtgctgctgctgctgctgctgctgctgctgctgctgc 4399
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 961 taactgctgctgacagcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgc 1020
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4400 taacttccaagcagagtgctgcttcaatgtgtccaggggtgtltaagcaggggttaactagg 4459
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1021 cctgtgtgcccgcctgcaagcacagcagctgcatctgcaagcagcga 1069
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4460 cctgtgtgcccgcctgtaaaaaactaaacacagcttgcgtgcaagttccca 4508
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13

US-10-105-299-806
; Sequence 806, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105, 299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197


```

; Prior Application removed - See file wrapper or Patent
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 806
; LENGTH: 1057
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-105-299-806

```

Query Match	5.98;	Score 97.4;	DB 6;	Length 1057;
Best Local Similarity	65.38;	Pred. NO. 4.2e-13;		
Matches 143; Conservative	0;	Mismatches 76;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 14
US-09-539-331D-15063
: Sequence 15063. Application US/09539331D
: GENERAL INFORMATION:
: APPLICANT: Sellhammer, Jeffrey J.
: APPLICANT: Deleagane, Angelo M.
: APPLICANT: Stuart, Susan G.
: APPLICANT: Stuve, Laura L.
: APPLICANT: Mullahy, Sara J.
: APPLICANT: Naughton, Rebecca E.
: TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
: FILE REFERENCE: PD-1022 CIP
: CURRENT APPLICATION NUMBER: US/09/539,331D
: CURRENT FILING DATE: 2000-03-30
: Prior Application removed - see file Wrapper or Palm
: NUMBER OF SEQ ID NOS: 40961
: SOFTWARE: PERL Program
: SEQ ID NO 15063
: LENGTH: 283
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID NO: hu00011877
US-09-539-331D-15063

```

Query Match 5.58; Score 92; DB 5; Length 283;
Best Local Similarity 64.68; Pred. No. 5.3e-12;
Matches 137; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY	858	g g g c t c c t g c g c t g c a c c a c c c t g t c t c c g g t c c c g c t c a a c t g t c g c g c g c g t g t c	917
	1	g g c a c a c t g c g c t g c a c a c c a c c a t g t c t c t g c a g c a g c t c a c t c g t t g t a c a a g a g t c	60
Db	918	c t t c a c g a g t g t g c t c t c t c t c g t g t g c g c c t g c c t g c c t g c t c t g t a c t c g c t c c a c g g	977
QY	918	c t t c a c g a g t g t g c t c t c t c t c t c g t g t g c g c c t g c c t g c c t g c t c t g t a c t c g c t c c a c g g	977
Db	61	a g c c a c g g g t g t c a g t a c c c c t c t t t c t g t t t a g t a g t a c c t c c a c c a a g g	120
QY	978	c t g c g t g a a g c t g a c c c a g c g t g c t a g a c a c g t c t c g c c g c c c t g t t g t c g c g t g c a a	1037
QY	121	t g c c t a a a t t g t g c a a g g t g t g t t a t a g a c c g g t t a a c a g g c t g t t g t c g c g t g t a a	180

```

QY      1038  gcacacgacacagcgtcatctgcaaaagcagcca 1065
          || ||||| | ||||| ||
Db      181  aaatcnaacacagtttctgctgcaaaagttccca 212

```

RESULT 15
US-09-540-210B-24047

```

Sequence 24047, Application US/09540210B
GENERAL INFORMATION:
APPLICANT: Selhamer, Jeffrey J.
APPLICANT: Deleogene, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
FILE REFERENCE: PD-1037 CIP
CURRENT APPLICATION NUMBER: US/09/540,210B
CURRENT FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: 08/972,899
PRIOR FILING DATE: November 18, 1997
PRIOR APPLICATION NUMBER: 08/395,244
PRIOR FILING DATE: February 27, 1995
PRIOR APPLICATION NUMBER: 08/722,922
PRIOR FILING DATE: September 27, 1996
PRIOR APPLICATION NUMBER: 60/005,526
PRIOR FILING DATE: September 29, 1995
PRIOR APPLICATION NUMBER: 08/824,029
PRIOR FILING DATE: March 25, 1997
PRIOR APPLICATION NUMBER: 60/014,010
PRIOR FILING DATE: March 25, 1996
PRIOR APPLICATION NUMBER: 08/826,847
PRIOR FILING DATE: April 10, 1997
PRIOR APPLICATION NUMBER: 60/015,533
PRIOR FILING DATE: April 10, 1996
PRIOR APPLICATION NUMBER: 08/993,555
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/023,308
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/862,178
PRIOR FILING DATE: May 22, 1997
PRIOR APPLICATION NUMBER: 60/018,217
PRIOR FILING DATE: May 23, 1996
PRIOR APPLICATION NUMBER: 08/881,589
PRIOR FILING DATE: June 24, 1997
PRIOR APPLICATION NUMBER: 60/021,275
PRIOR FILING DATE: June 25, 1996
PRIOR APPLICATION NUMBER: 08/903,802
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/023,308
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/905,881
PRIOR FILING DATE: August 1, 1997
PRIOR APPLICATION NUMBER: 60/025,204
PRIOR FILING DATE: August 1, 1996
PRIOR APPLICATION NUMBER: 08/903,471
PRIOR FILING DATE: July 30, 1997
PRIOR APPLICATION NUMBER: 60/025,478
PRIOR FILING DATE: July 31, 1996
PRIOR APPLICATION NUMBER: 08/903,556
PRIOR FILING DATE: July 31, 1997
PRIOR APPLICATION NUMBER: 60/025,217
PRIOR FILING DATE: August 22, 1996
PRIOR APPLICATION NUMBER: 08/937,142
PRIOR FILING DATE: September 23, 1997
PRIOR APPLICATION NUMBER: 60/026,598
PRIOR FILING DATE: September 24, 1996
PRIOR APPLICATION NUMBER: 08/960,746
PRIOR FILING DATE: October 29, 1997
PRIOR APPLICATION NUMBER: 60/030,144
PRIOR FILING DATE: October 30, 1996
PRIOR APPLICATION NUMBER: 08/926,847
PRIOR FILING DATE: April 10, 1997

```


PRIOR APPLICATION NUMBER: 60/015,533
PRIOR FILING DATE: April 10, 1996
PRIOR APPLICATION NUMBER: 08/755,524
PRIOR FILING DATE: November 22, 1996
PRIOR APPLICATION NUMBER: 60/007,495
PRIOR FILING DATE: November 22, 1995
PRIOR APPLICATION NUMBER: 09/021,031
PRIOR FILING DATE: February 10, 1998
PRIOR APPLICATION NUMBER: 60/039,325
PRIOR FILING DATE: February 13, 1997
PRIOR APPLICATION NUMBER: 09/035,172
PRIOR FILING DATE: March 4, 1998
PRIOR APPLICATION NUMBER: 60/040,431
PRIOR FILING DATE: March 5, 1997
PRIOR APPLICATION NUMBER: 09/041,894
PRIOR FILING DATE: March 12, 1998
PRIOR APPLICATION NUMBER: 60/040,199
PRIOR FILING DATE: March 14, 1997
PRIOR APPLICATION NUMBER: 09/050,817
PRIOR FILING DATE: March 30, 1998
PRIOR APPLICATION NUMBER: 60/043,792
PRIOR FILING DATE: April 11, 1997
PRIOR APPLICATION NUMBER: 09/074,999
PRIOR FILING DATE: May 8, 1998
PRIOR APPLICATION NUMBER: 60/048,431
PRIOR FILING DATE: May 29, 1997
PRIOR APPLICATION NUMBER: 09/107,592
PRIOR FILING DATE: June 30, 1998
PRIOR APPLICATION NUMBER: 60/052,751
PRIOR FILING DATE: July 1, 1997
PRIOR APPLICATION NUMBER: 09/094,079
PRIOR FILING DATE: June 9, 1998
PRIOR APPLICATION NUMBER: 60/049,975
PRIOR FILING DATE: June 13, 1997
NUMBER OF SEQ ID NOS: 35654
SOFTWARE: PERL Program
SEQ ID NO 24047
LENGTH: 147
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID NO: hu00410209
LOCATION: 33, 57
NAME/KEY: unsure
OTHER INFORMATION: a, t, c, g, or other
US-09-540-210B-24047

Query Match 4.6%; Score 76.8; DB 5; Length 147;
Best Local Similarity 69.9%; Pred. No. 1.2e-08;
Matches 102; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 661 gagctgacacagactctctgctgacagagcctgtggaagttaaatcaagagagt 720
|| ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 gatttggcctcagccctcactcagtgtaggncctgtgcaagtccaatttaagmgctc 61
QY 721 gcataccccgcgacttgccttcctgctgctgcaacacagagtgctgtcagcc 780
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 acctaccacaaagcctcctgacacacagatgactcgcacaaagcagtgcttgcgcgc 121
QY 781 caagactctgcaactatgacagtg 806
||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 cagaagctgactatgactgtgactgt 147

Search completed: May 11, 2002, 06:20:51
Job time: 13101 sec

A:Cross-references: EMBL:M25864
 R:Bonthron, D.; Orkin, S.H.
 A:Title: The human von Willebrand factor gene. Structure of the 5' region.
 A:Reference number: S02377; MUID:86111704
 A:Accession: S02377
 A:Molecule type: DNA
 A:Residues: 1-177 <R02>
 A:Cross-references: EMBL:X06828
 R:Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Lester-Mancuso, T.L.; Le Beau, M.M.; Sorio
 R:Biochemistry 30, 253-269, 1991
 A:Title: Human von Willebrand factor gene and pseudogene: structural analysis and differ
 A:Reference number: A37139; MUID:91105089
 A:Accession: A37139
 A:Molecule type: DNA
 A:Residues: 390-1947 <MAD>
 A:Cross-references: GB:M60675; NID:9340357; PIDN:AAA61295.1; PID:9553810
 A:Note: the authors translated the codon CGC for residue 156 as Gln
 R:Collins, C.J.; Underdahl, J.P.; Levene, R.B.; Ravera, C.P.; Morin, M.J.; Dombalagian,
 Proc. Natl. Acad. Sci. U.S.A. 84, 4393-4397, 1987
 A:Title: Molecular cloning of the human gene for von Willebrand factor and identificatio
 A:Reference number: S23676; MUID:87260814
 A:Accession: S23676
 A:Molecule type: DNA
 A:Residues: 2731-2813 <COL>
 A:Cross-references: EMBL:M16945
 R:Bonthron, D.; Orr, E.C.; Mitsuoka, L.M.; Ginsburg, D.; Handin, R.I.; Orkin, S.H.
 Nucleic Acids Res. 14, 7125-7127, 1986
 A:Title: Nucleotide sequence of pre-pro-von Willebrand factor cDNA.
 A:Reference number: A25298; MUID:87016349
 A:Accession: A25298
 A:Molecule type: mRNA
 A:Residues: 1-470, 'V', 472-2813 <BON>
 A:Cross-references: EMBL:X04385
 R:Verweij, C.L.; Diegaarde, P.J.; Hart, M.; Pannekoek, H.
 EMBO J. 5, 1839-1847, 1986
 A:Title: Full-length von Willebrand factor (vWF) cDNA encodes a highly repetitive protei
 A:Reference number: A91044; MUID:87004550
 A:Accession: A25469
 A:Molecule type: mRNA
 A:Residues: 1-470, 'V', 472-483, 'R', 485-1022, 'K', 1024-1025, 'E', 1027-1400 <VER>
 A:Cross-references: EMBL:X04146
 A:Note: this sequence has been revised in reference A91056
 R:Verweij, C.L.; Diegaarde, P.J.; Hart, M.; Pannekoek, H.
 EMBO J. 5, 3074, 1986
 A:Reference number: A91056
 A:Accession: A25366
 A:Molecule type: mRNA
 A:Residues: 1021-1030 <VE2>
 A:Note: this is a revision to the sequence from reference A91044
 R:Shelton-Inloes, B.B.; Broze Jr., G.J.; Miletich, J.P.; Sadler, J.E.
 Biochem. Biophys. Res. Commun. 144, 657-665, 1987
 A:Title: Evolution of human von Willebrand factor: cDNA sequence polymorphisms, repeated
 A:Reference number: S23618; MUID:87213253
 A:Accession: S23618
 A:Molecule type: mRNA
 A:Residues: 1-120 <SH2>
 A:Cross-references: EMBL:M17588; NID:9799330; PIDN:AAA65940.1; PID:9340316
 A:Accession: S23645
 A:Molecule type: protein
 A:Residues: 23-56 <SH3>
 R:Sadler, J.E.; Shelton-Inloes, B.B.; Sorace, J.M.; Harlan, J.M.; Tilani, K.; Davie, E.W.
 Proc. Natl. Acad. Sci. U.S.A. 82, 6394-6398, 1985
 A:Title: Cloning and characterization of two cDNAs coding for human von Willebrand facto
 A:Reference number: A94060; MUID:86016708
 A:Accession: A94060
 A:Molecule type: mRNA
 A:Residues: 'WA', 739, 'C', 744-769, 'H', 771-788, 'A', 790-803, 'S', 805-873, 1289-1471, 'D', 1473-
 A:Note: The authors translated the codon TCG for residue 2168 as Cys
 R:Shelton-Inloes, B.B.; Tilani, K.; Sadler, J.E.
 Biochemistry 25, 3164-3171, 1986
 A:Title: cDNA sequences for human von Willebrand factor reveal five types of repeated do
 A:Reference number: A90504; MUID:86269894

A:Accession: A90504
 A:Molecule type: mRNA
 A:Residues: 781-788, 'A', 790-1424 <SHE>
 A:Note: 852-Gln, 857-Asp, and 1381-Thr were also found
 R:Ginsburg, D.; Handin, R.I.; Bonthron, D.T.; Doulton, T.A.; Bruns, G.A.P.; Latt, S.A.
 Science 228, 1401-1406, 1985
 A:Title: Human von Willebrand factor (vWF): isolation of complementary DNA (cDNA) cto
 A:Reference number: A44178; MUID:85244588
 A:Accession: A44178
 A:Molecule type: mRNA
 A:Residues: 2621-2813 <GIN>
 A:Cross-references: EMBL:X02672; NID:9340308; PIDN:AAA61293.1; PID:9340309
 R:Verweij, C.L.; de Vries, C.J.M.; Distel, B.; van Zonneveld, A.J.; van Kessel, A.G.;
 Nucleic Acids Res. 13, 4699-4717, 1985
 A:Title: Construction of cDNA coding for human von Willebrand factor using antibody p
 A:Reference number: S07363; MUID:85266603
 A:Accession: S07363
 A:Molecule type: mRNA
 A:Residues: 2731-2813 <VE3>
 A:Cross-references: EMBL:X02672; NID:937939; PIDN:CAA26503.1; PID:937940
 R:Lynch, D.C.; Zimmerman, T.S.; Collins, C.J.; Brown, M.; Morin, M.J.; Ling, E.H.; Li
 Cell 41, 49-56, 1985
 A:Title: Molecular cloning of cDNA for human von Willebrand factor: authentication by
 A:Reference number: S23678; MUID:85201687
 A:Accession: S23678
 A:Molecule type: mRNA
 A:Residues: 2731-2813 <LIR>
 A:Cross-references: EMBL:X03028
 R:Tilani, K.; Kumar, S.; Takio, K.; Ericsson, L.H.; Wade, R.D.; Ashida, K.; Walsh, K.
 Biochemistry 25, 3171-3184, 1986
 A:Title: Amino acid sequences of human von Willebrand factor.
 A:Reference number: A90505; MUID:86269895
 A:Accession: A90505
 A:Molecule type: protein
 A:Residues: 764-788, 'A', 790-1471, 'D', 1473-2813 <TT>
 A:Note: 789-Thr was also found
 R:Chopex, M.W.; Gitma, J.P.; Fujikawa, K.; Davie, E.W.; Tilani, K.
 Biochemistry 25, 3146-3155, 1986
 A:Title: Human von Willebrand factor: a multivalent protein composed of identical sub
 A:Reference number: A23464; MUID:86269892
 A:Accession: A23464
 A:Molecule type: protein
 A:Residues: 764-773, 2803-2813 <CHO>
 R:Dent, J.A.; Berkowitz, S.D.; Ware, J.; Kasper, C.K.; Ruggeri, Z.M.
 Proc. Natl. Acad. Sci. U.S.A. 87, 6306-6310, 1990
 A:Title: Identification of a cleavage site directing the immunochemical detection of
 A:Reference number: A36013; MUID:90349604
 A:Accession: A36013
 A:Molecule type: protein
 A:Residues: 1606-1617 <DEN>
 R:Ray, P.J.; Kawai, Y.; Wagner, D.D.; Ginsburg, D.; Bonthron, D.; Ohlsson-Wilhelm, B.
 Science 232, 995-998, 1986
 A:Title: Propolypeptide of von Willebrand factor circulates in blood and is identical
 A:Reference number: A60913; MUID:86208144
 A:Accession: A60913
 A:Molecule type: protein
 A:Residues: 576-590 <RAY>
 C:Genetics:
 A:Gene: GDB:VWF
 A:Cross-references: GDB:119125; OMIM:193400
 A:Map position: 12p13.3-12p13.2
 A:Anticodon: 19/1, 74/1, 108/2, 178/1, 218/3, 292/1, 333/1, 370/2, 386/1, 431/3, 478/1,
 5/2, 1724/1, 1771/1, 1819/1, 1874/1, 1888/3, 1948/1, 2021/3, 2086/1, 2200/1, 2266/3, v
 C:Superfamily: von Willebrand factor, von Willebrand factor type A repeat homology, v
 C:Keywords: blood coagulation; cell binding; connective tissue; disulfide bond; dupli
 F:1-22/Domain: signal sequence; status predicted <SIG>
 F:23-763/Product: von Willebrand antigen II #status predicted <NAL>
 F:734-386/Domain: type D repeat 1 <DD1>
 F:387-745/Domain: type D repeat 2 <DD2>
 F:698-700/Region: cell attachment (R-G-D) motif
 F:764-2813/Product: von Willebrand factor #status predicted <MA2>
 F:784-865/Domain: D, <DDD>
 F:788-833, 2216-2261/Region: duplication

F:826-853,2400-2515,2544-2662/Region: duplication
 F:842-1130,1934-2203/Region: duplication
 F:866-1241/Domain: type D repeat 3 <DD3>
 F:1275-1443/Domain: von Willebrand factor type A repeat homology <VWA1>
 F:1496-1654/Domain: von Willebrand factor type A repeat homology <VWA2>
 F:1689-1854/Domain: von Willebrand factor type A repeat homology <VWA3>
 F:1947-2295/Domain: type D repeat 4 <DD4>
 F:2296-2330/Domain: type B repeat 1 <VB1>
 F:2340-2365/Domain: type B repeat 2 <VB2>
 F:2375-2399/Domain: type B repeat 3 <VB3>
 F:2430-2497/Domain: von Willebrand factor type C repeat homology <VWC1>
 F:2507-2509/Region: cell attachment (R-G-D) motif
 F:2581-2647/Domain: von Willebrand factor type C repeat homology <VWC2>
 F:557,1231,1515,1574,2223,2290,2357,2400,2546,2585,2790/Binding site: carbohydrate (Asn)
 F:1147/Binding site: carbohydrate (Asn) (covalent) #status atypical
 F:1246,1255,1256,1468,1477,1487,1679,2298/Binding site: carbohydrate (Thr) (covalent) #s
 F:1263,1486/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 6.7%: Score 110.5; DB 1; Length 2813;
 Best Local Similarity 21.0%: Pred. No. 3.3;
 Matches 56; Conservative 30; Mismatches 90; Indels 91; Gaps 13;

OY 77 TPARKQDVTTHMISFGSPSSSSSTSDQRLDHMAFPVDAQSPRAVRIOPKV 136
 DB 727 TMCYCEDEGMH--CTMSGVPSGLPDVAVLSSP---LSH-----RKRSLSCRPV 772
 OY 137 H-----CQPLDK-----GPAVPELDKHFLLCEACGCKC--- 167
 DB 773 KLVCPADNLRABGLCTCTQCNVDECMSCMGVSGICLCPGMVHRNRCVALERCPCHQ 832
 OY 168 -KECASPTLPSCWVCNOECCLCSAQTIVNCTCMLQVGFYHCTNEDEGSCADHPCC-- 224
 DB 833 GKRYAGCEIVK-----ICGNTVC-----RDKMCTHVDCA 865
 OY 225 SCGRSMCCARMSFGALSVYLPCLLYL-----PATGCVKLAQRYDLRRPGCC 275
 DB 866 TCSTIGMAHYLTFDG-LKYLFPGEQVYLVODYCGSNPFFRILVGNKG---CSHPVSK 921
 OY 276 KHTNSVYKAA-----SCDAKTSRPDK 297
 DB 922 KKRVTILVEGGELELFDGQVNNKRRMK 948

RESULT 6
 T14764
 hypothetical protein DKFZP434H204.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T14764
 R:Hamblitt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 Submitted to the Protein Sequence Database, August 1999
 A:Reference number: 218181
 A:Accession: T14764
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-898 <WAM>
 A:Cross-references: EMBL:AL110226
 A:Experimental source: adult testis; clone DKFZP434H204
 C:Genetics:
 A:Note: DKFZP434H204.1

Query Match 6.6%: Score 109; DB 2; Length 898;
 Best Local Similarity 23.5%: Pred. No. 1.4;
 Matches 80; Conservative 18; Mismatches 125; Indels 118; Gaps 22;

OY 5 IPQSAPLTPNSVWOPLDLSRMSHSLQHLPLTLPIDQVKSHEVDYID---NPFLAT 61
 DB 600 VPETPLAPSLAEAGPPAD-----PLVVR-----NNSWAGNMWSECSITTCGLGAV 644
 OY 62 TGPKRTGCG---APELAPTPA-RCDDQVTHMISFGSPSSSSSTSDQRLDHMA 115

DB 645 WRPVRCSSGDEDCAPRGPQPARRCHLRPCATW--HSGNMSKCSRSRGSSVRDQ-- 700
 OY 116 APPVADQSPRAVRIOPKVHQCQPLDLKGPVAPPE-----LDKHFLLC-EAC 162
 DB 701 ----CVDRDLRPLR-----PFHCQ-----GPAKPRANRCSGADPCLSYTSSMRECEAC 748
 OY 163 G-----KCK-----CKEASPTLPSC-----WYCNQECCLCSAQTIVNCTCMLY 203
 DB 749 GGGEDQRLVTCPEFGCEFLARNTTRPCNTTRPCNTQVWVGPWGQSA-----PCGGV 801
 OY 204 QGIFVHCTN-----EDDEGSC-----ADHPCS-----CARW 235
 DB 802 QRRLVKCVNTQTGLPEEDSDQCHENAPSSRPGCTGEDCEPYEPRCERDLSTFGCEYL 861
 OY 236 SPMGALSVLP-----CLCYLPATGCVKLAQRYDLRLR 270
 DB 862 RLHGRCQ--LPTITQCCRSCSPSHGA---PSRCHQRYAR 897

RESULT 7
 A35672
 crumbs protein - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 21-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 11-Jan-2000
 C:Accession: A35672
 R:Repass, U.; Theres, C.; Knust, E.
 Cell 61, 787-799, 1990
 A:File: crumbs encodes an EGF-like protein expressed on apical membranes of Drosophi
 A:Reference number: A35672; MUID:90283104
 A:Accession: A35672
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2139 <TEP>
 A:Cross-references: GB:M3753
 A:Note: the authors translated the codon GGC for residue 1928 as Cys, and TAT for res
 C:Genetics:
 A:Gene: flybase:cb
 A:Cross-references: Flybase:FBgn0000368
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: transmembrane protein
 F:352-385/Domain: EGF homology <EGX1>
 F:392-424/Domain: EGF homology <EGF1>
 F:691-722/Domain: EGF homology <EGF>
 F:767-799/Domain: EGF homology <EGF3>
 F:1878-1914/Domain: EGF homology <EGX2>

Query Match 6.5%: Score 108; DB 2; Length 2139;
 Best Local Similarity 25.8%: Pred. No. 3.8;
 Matches 63; Conservative 15; Mismatches 82; Indels 84; Gaps 15;

OY 68 RCGAPELAPTPA-RCDDQ-----YTHMISFGSPSSSSSTSDQ----- 109
 DB 1658 KGHESWDPPTAFSTDDIDQAFQSLATSTOVYLGCMPEBQARGSTLSAQGSGFKCV 1717
 OY 110 ----RLDHMAWP-PVADQSPRAVRIOPKVHQCQPLDLKGPVAPPELDKHFLLCEACGCKC 165
 DB 1718 GEARLDLILPYFSMAELYSRTNVSVOOKA-----QFRLNATRPEEG-----CLICFOS 1766
 OY 166 KCKE---CASPTLPSCWVC-----NOECCLCSAQTIVNCTCMLYOG 205
 DB 1767 DCKNDGFCQSPDEYAC-TCQPGEGDCCGTDDDECLNTCL-----NNGTCTGNYAA 1818
 OY 206 IF-----YHCTNEDEGSCADHPCCSRSC-----CARMSFGALSVLP 246
 DB 1819 FFCQCPGFGEGCHCEQNIDE---CADQPCN-NGGNCITDLIASVYCDCE-DMGPOCDVLK 1874
 OY 247 CLIC 250
 DB 1875 QMTC 1878

RESULT 8

A:introns: 23/3; 71/3; 121/3; 183/1; 1083/3; 1141/3; 1408/1; 1449/1

Query Match 6.4%; Score 106.5; DB 2; Length 1513;
Best Local Similarity 21.0%; Pred. No. 3.6;
Matches 73; Conservative 38; Mismatches 123; Indels 113; Gaps 20;

OY 2 EPPIPQASAPLTPNSVWVQPLDLSRMSHRLQHPILTLIPDOYKTSVENDYIDNPISALT 61
DB 1181 QPTVQOAPAP-----OLVQOPI-----YQTVQAPLSPVPVPAVPS-----APSOA-- 1220
OY 62 TGPRTTRGGA-PELAPTRACDDVT-----HHWISFSGRSSVSSSSSTSSDQRL 111
DB 1221 SCPTITTFANAPCEPIV---CEQSCNTQCEKFPQDHGSCVNSQCCACATQTPAVQ-- 1275
OY 112 LDHMAPPP-----VADASPAVRIQPKV-----HCOPL---DLK 144
DB 1276 --AAPASPCOPCOPACEPVCAIQQAQP--VRIQINLATASSVLQASDADQPCBOQSCV 1330
OY 145 GPVAVPELDKHFLLCEACGKCKECSAP-----RTLPP-----SCWVCNOEC 186
DB 1331 QECQSTTLNQAATCQACQSCICQSCAPLGTSAVPMQITIPVPAFAVASTOLCAFKC 1390
OY 187 LCSAOTLVNNTGCMIVOGIFHYHNTNDEDEGSCADHPSC-----SRNCCARMSFMGLS 242
DB 1391 ISDCGLCKSKNSPQC-10GCDASC-----OQLCGTAPTPAVPLTVNYKC----- 1433
OY 243 VVLPC-----LCLYLPATGCVKLAQNGYDLRRPGCRCKHTNSVTC 284
DB 1434 -NLPCDOCTOOCVHQAPTCAPACASACE-AQCPIVASCBDACQTVCK 1478

RESULT 12

142215
zonadhesin - mouse
N:Alternate names: sperm-specific membrane protein
C:Species: Mus musculus (house mouse)
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T42215
R:Gao, Z.; Garbers, D.L.
J: Biol. Chem. 273, 3415-3421, 1998
A:Title: Species diversity in the structure of zonadhesin, a sperm-specific membrane protein
A:Reference number: 222080; MUID:98123114
A:Accession: T42215
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-5376 <GMO>
A:Cross-references: EMBL:U97068; NID:93327420; PID:93327421; PIDN:AAC26680.1
C:Genetics:
A:Gene: Zan
A:Map position: 5
A:Function:
A:Description: functions in multiple cell adhesion processes
A:Note: found exclusively on the apical region of the sperm head
C:Keywords: cell adhesion

Query Match 6.3%; Score 105; DB 2; Length 5376;
Best Local Similarity 22.3%; Pred. No. 15;
Matches 65; Conservative 23; Mismatches 108; Indels 96; Gaps 15;

OY 54 DNPSTLTTGPKRTGCAPELAPPARCDDVTHHMTSFGSRSSVSSSSSTSSDORLID 113
DB 4079 DKCVLRIGCGCKDTGGILPAGRT-----WIS-----SDCTKSCSGMGIIQCRD 4123
OY 114 HMAPPPVADQASPRAVRIQPKVHCPDLKGPVAPPDLKHFLC-EAC-----GK 164
DB 4124 FGCPPTCKESNDSSRT-----CAKIPQCPA-----HSHTNCLPACNSRCTDLIGH 4172
OY 165 CKKCKEASRRLPSCVNCQDELTCAQTLVNYGCMKLVGICITVHCNENDEDE----- 217
DB 4173 C---EGTSPK-VPS--PCKESGLCOPGTIVNHKCVLQI-----HGCKDAQGGFVPAGK 4221

OY 218 -----SCADHPGCSNSNCCARMSFEMGALSVLPCL 248
DB 4222 TWISRGCTQSCACGAVOCHNFTPTGTGTCQNSSCSKITYOCFMSQITTCPLSPFS- 4280
OY 249 LCVLPATGCVKLAQNGYDLRRPGRC-----KHTNSYI-----CKAASGD 289
DB 4281 -CFDPBELCGGASPRAPSTCRG-GCVCEADYVLRDXKVLITQCGCKDAQDG 4330

RESULT 13

A29072
pulmonary surfactant protein SP 18 precursor - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Aug-1999
C:Accession: B29072; A29072
R:Hawgood, S.; Benson, B.J.; Schilling, J.; Damm, D.; Clements, J.A.; White, R.T.
Proc. Natl. Acad. Sci. U.S.A. 84, 66-70, 1987
A:Title: Nucleotide and amino acid sequences of pulmonary surfactant protein SP 18 an
A:Reference number: A29072; MUID:87092398
A:Accession: B29072
A:Molecule type: mRNA
A:Residues: 1-363 <HAW>
A:Cross-references: GB:M15170; NID:9164077; PIDN:AAA30893.1; PID:9164078
A:Accession: A29072
A:Molecule type: protein
A:Residues: 182-210 <H2>
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology
F:1-14/Domain: signal sequence #status predicted <SIG>
F:15-180/Domain: propeptide #status predicted <RP>
F:54-146/Domain: saposin repeat homology <SAP1>
F:180-267/Domain: saposin repeat homology <SAP2>
F:181-363/Product: pulmonary surfactant protein SP 18 #status experimental <MAT>
F:273-358/Domain: saposin repeat homology <SAP3>

Query Match 6.3%; Score 104; DB 2; Length 363;

Best Local Similarity 22.8%; Pred. No. 1.4;
Matches 54; Conservative 24; Mismatches 83; Indels 76; Gaps 11;

OY 130 RIQPKVYHCOPLDKGPVAP-----PELDKHFLLCEACGCKR-----CKECA 171
DB 123 QINPKITL-CKHGLGCKPGLPEPQESLSDPLDK-LIIPELGALVYGPHTDLSQQ 180
OY 172 SPRTLPPSCWVNOEC-----LCSAOTLVNNTGCMV----- 201
DB 181 LPITLPYCMCLRLIKRIQAMIRKGVLAATVGGVCHVPLVVGICQCLGERITYVLLDA 240
OY 202 -----LVGITYHCTNEDDEG-SCADHPSCS-RSNCCARMSFEMGA----- 240
DB 241 ILGRMLPQLVGLVLRCSHEDSAGPALASLPSEWSFQESKCOLCMFVTTQAGNHSQATP 300
OY 241 ISVLPCLLCLYLPATGCVKLAQNGYDLR--RPGCRCKHTNSVTCASGDAKTSRP 295
DB 301 QAIRACGLSSWLDHQKCEQFVEQHMPRLQTLASGGRDAHTT---COALGACRTTFSP 354

RESULT 14

T27283
hypothetical protein Y64G10A.f - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27283
R:Almscough, R.
Submitted to the EMBL Data Library, September 1999
A:Reference number: Z20336
A:Accession: T27283
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1620 <WIL>
A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CA54471.1; CBSP:Y64G10A.f
A:Experimental source: clone Y64G10A
C:Genetics:
A:Gene: CBSP:Y64G10A.f

Query Match	6.3%;	Score 104;	DB 2;	Length 1620;
Best local similarity	26.5%;	Prod no 60;		

Matches	41; Conservative	18; Mismatches	58; Indels	38; Gaps	10
---------	------------------	----------------	------------	----------	----

158 LCEA-CGCKCKE-----CASPTLPSCWVCNQC-----LCSAOTLVNYGTGM 200

1371 MCDATSGCFCFKPGHSGSDCKSGCYOGRFPD---CNDICSPENGVCYDSSST---GSCV 1433

201 CLVGGTFYHC---TNEDEFGSCADHDCSCSBSNCGA DVGKMGAT AUNR DQQT GVV----- 689

```

1424  C P B G V I G M K C E T A C C S D B E C D M C N V R C O N C A C C T T T T
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
      -----
      257

```

-----LITGQCRCRLPFTGMT-C 1476

256 VALLAQK-YDRLRPGCRCKHNSVICAKASGDAK 291

1477 NQVPEGRFGAGCKEKCRCANGH---CNASSGECK 1508

SQRT 15

7057

Species: *Cavia porcellus* (guinea pig)

Accession: B37057
Date: 15-Feb-1991 #sequence_revision 13-Sep-1991 #text_change 20-Aug-1999

Sheppard, D.; Rozzo, C.; Starr, L.; Quaranta, V.; Erle, D.J.; Pytela, R. *Biol. Chem.* 265, 11502-11507, 1990

Title: Complete amino acid sequence of a novel integrin beta subunit (beta6)
Reference number: A27057, with 0020755

Accession: B37057
Accession number: A37057; MUID:9030/659

status: preliminary
molecule type: mRNA

Residues: 1-577 <SHE>
Cross-references: GB:M35197; GB:J05522; NID:d191277; BDN:AAA370A3 1; STD:-EFF

Note: the authors translated the codon AAA for residue 88 as Asn, AAC for res

ly

Keywords: cell adhesion; cytoskeleton; transmembrane protein

Query Match	6.28; Score 102.5; DB 2; Length 577;
Best Local Similarity	27.68; Pred. No. 2 8.

Matches	45;	Conservative	12;	Mismatches	37;	Indels	69;	Gaps	13;
---------	-----	--------------	-----	------------	-----	--------	-----	------	-----

157 ILCEACGKCKCKECCASPTLPSCW---CNGECLCSAQTLVN-----YGTMC--LV 203

409 LCGDNGDCEGEC---VCRSGWTGEYCN--CTTSTDTCISEGTLCSGRGDPYVGKCY 462

204 061FYHCTNEDDEG-----SCADHPCSCSRNCCAPBSEWCAI SVUJ DOU LOU VUJ DAME JEC

```

      ||| - :|||:- |  

-----CTNPGA SCBPTC EBCD PMS SD- DONSERDC -----  


```

-----CIECHLSDC 498

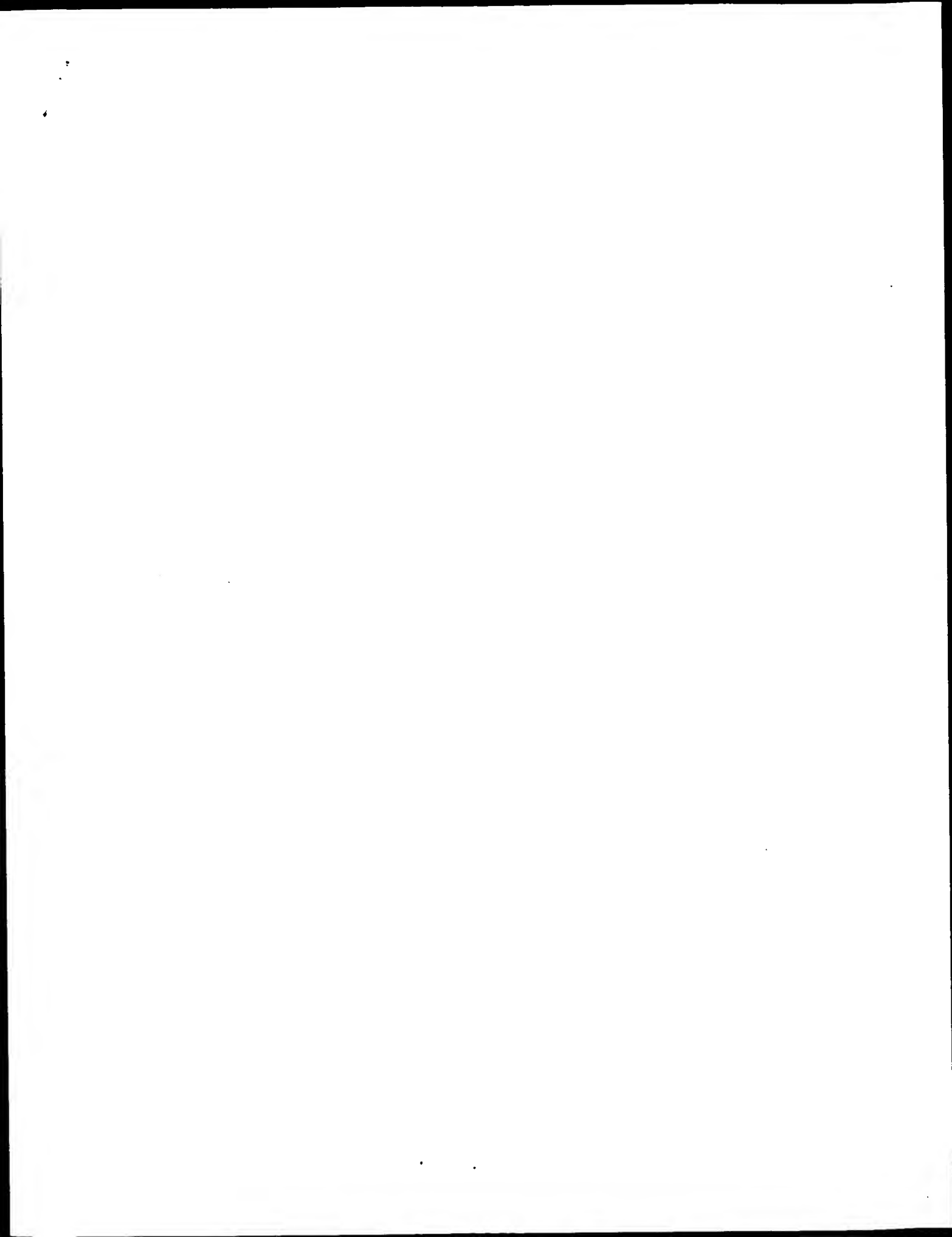
237 CVALAQSGIDKRNPG--C--RCKHNSVICAA--SGDAKTS 293

499 -----QPGEECVDKCLAGVTISKADPSKDSYS 528

Arch completed: May 8, 2002, 10:46:40
Time: 53 sec

המחיר: 25 ש"ח

100



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:45:48 ; Search time 8.33 seconds
(without alignments)
640.295 Million cell updates/sec

Title: US-10-082-902-2

Perfect score: 1657
1 MEPPPOSAPLPTNSVMWOP.....SVICKASGAKTSRPDKPF 299

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 81999 seqs, 17638306 residues

Total number of hits satisfying chosen parameters: 81999

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: Pending_Patents_AA_New:*
2: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1657	100.0	299	1 PCT-US02-08781-8	Sequence 8, Appl1
2	1657	100.0	299	6 US-10-082-902-2	Sequence 2, Appl1
3	1621.5	97.9	375	1 PCT-US02-09921-1088	Sequence 1088, Ap
4	1554.5	93.8	300	6 US-10-082-902-5	Sequence 5, Appl1
5	641.5	38.7	315	6 US-10-082-902-3	Sequence 3, Appl1
6	634.5	38.3	315	6 US-10-082-902-6	Sequence 6, Appl1
7	528	31.9	139	6 US-10-082-902-4	Sequence 4, Appl1
8	237	14.3	100	6 US-10-082-902-7	Sequence 7, Appl1
9	129	7.8	132	6 US-10-123-155-255	Sequence 255, App
10	121	7.3	299	6 US-10-123-155-39	Sequence 39, Appl
11	116	7.0	1162	6 US-10-123-155-185	Sequence 185, App
12	114.5	6.9	3323	6 US-10-123-155-113	Sequence 113, App
13	114	6.8	2477	6 US-10-123-155-331	Sequence 331, App
14	112	6.8	2290	6 US-10-123-155-323	Sequence 323, App
15	111.5	6.7	633	1 PCT-US02-09944-533	Sequence 533, App
16	111.5	6.7	738	1 PCT-US02-09944-712	Sequence 712, App
17	111.5	6.7	3552	6 US-10-123-155-339	Sequence 339, App
18	111	6.7	2163	6 US-10-123-155-127	Sequence 127, App
19	110.5	6.6	1871	6 US-10-063-545-91	Sequence 91, Appl
20	110	6.6	2447	6 US-10-123-155-291	Sequence 291, App
21	110	6.6	2542	6 US-10-123-155-95	Sequence 95, Appl
22	109.5	6.6	771	6 US-10-123-155-495	Sequence 495, App
23	109.5	6.6	1701	6 US-10-063-545-73	Sequence 73, Appl
24	109	6.6	469	1 PCT-US02-09785-980	Sequence 980, App
25	109	6.6	494	1 PCT-US02-09785-762	Sequence 762, App
26	108.5	6.5	422	6 US-10-063-545-65	Sequence 65, Appl

27	108.5	6.5	2849	6 US-10-123-155-371	Sequence 371, App
28	108.5	6.5	3479	6 US-10-123-155-123	Sequence 123, App
29	108	6.5	1312	6 US-10-123-155-397	Sequence 397, App
30	108	6.5	1676	6 US-10-063-545-53	Sequence 53, Appl
31	108	6.5	2762	6 US-10-123-155-13	Sequence 13, Appl
32	107.5	6.5	1857	6 US-10-123-155-365	Sequence 365, App
33	107.5	6.5	2651	6 US-10-063-545-23	Sequence 23, Appl
34	107.5	6.5	3884	6 US-10-123-155-145	Sequence 145, App
35	107.5	6.5	3907	6 US-10-029-2174-24	Sequence 24, Appl
36	107	6.5	804	6 US-10-123-155-149	Sequence 149, Appl
37	107	6.5	1234	6 US-10-063-545-63	Sequence 63, Appl
38	107	6.5	1837	6 US-10-123-155-243	Sequence 243, App
39	107	6.5	2236	6 US-10-123-155-381	Sequence 381, App
40	107	6.5	3819	6 US-10-123-155-405	Sequence 405, App
41	106.5	6.4	1971	6 US-10-123-155-139	Sequence 139, App
42	106.5	6.4	2768	6 US-10-063-545-15	Sequence 15, Appl
43	106	6.4	1523	6 US-10-123-155-429	Sequence 429, App
44	106	6.4	1570	6 US-10-123-155-455	Sequence 455, App
45	106	6.4	1576	6 US-10-123-155-301	Sequence 301, App

ALIGNMENTS

RESULT 1
PCT-US02-08781-8
: Sequence 8, Application PC/TUS0208781
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc
: APPLICANT: Yang, Tom Y.
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle
: APPLICANT: Asundi, Vinod
: APPLICANT: Liu, Chenghua
: APPLICANT: Wehrman, Tom
: APPLICANT: Ren, Felvan
: APPLICANT: Drmanac, Radote, T.
: TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
: FILE REFERENCE: 21272-006-061 CIP
: CURRENT APPLICATION NUMBER: PCT/US02/08781
: PRIOR FILING DATE: 2002-03-20
: PRIOR APPLICATION NUMBER: 09/814,354
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 8
: LENGTH: 299
: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-US02-08781-8

Query Match 100.0%; Score 1657; DB 1; Length 299;

Best Local Similarity 100.0%; Pred. No. 4,2e+115; Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MEPPIPOSAPALTNSVMWOP	PLDSSKMSHSRLQHLPTLTPIDQVKTSHVENDYIDNPSLAL	60
Db	1	MEPPIPOSAPALTNSVMWOP	PLDSSRMSHSRLQHLPTLTPIDQVKTSHVENDYIDNPSLAL	60
QY	61	TTGPKRTGGAPELATTPARCQDQVTHHMSFSGSPSSVSSSTSDQRLDHAAPPV		12
Db	61	TTGPKRTGGAPELATTPARCQDQVTHHMSFSGSPSSVSSSTSDQRLDHAAPPV		12
QY	121	ADQASPAVRIQPKVVHCGPRLDKGPVAPPELDHFLICACGCKCKCKEASPRLLPSCW		18
Db	121	ADQASPAVRIQPKVVHCGPRLDKGPVAPPELDHFLICACGCKCKCKEASPRLLPSCW		18
QY	181	VCHOEPLCSAQTLVNYGTOMCLYOGITFYHCNEDDESCADHPSCSRNCCARMSFKA		24
Db	181	VCHOEPLCSAQTLVNYGTOMCLYOGITFYHCNEDDESCADHPSCSRNCCARMSFKA		24
QY	241	LSVVLPCLLCYLPATGCVKLAQRYDLRLRPPGCRCKHTNSVICKASGDAKTSRPDKPF		299

Db 241 LSVYLPCLLCYLPATGCVKLAQRGIDRLRRPGCKCKHTNSVICKAASGDAKTSRPDKPF 299

RESULT 2
US-10-082-902-2
Sequence 2, Application US/10082902
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
APPLICANT: Montgomery, Julie R.
TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
FILE REFERENCE: 28110/35878
CURRENT APPLICATION NUMBER: US/10/082,902
PRIOR FILING DATE: 2002-02-26
PRIOR APPLICATION NUMBER: US/09/370,398
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 299
TYPE: PRT
ORGANISM: Homo sapiens
US-10-082-902-2

Query Match 100.0%; Score 1657; DB 6; Length 299;
Best Local Similarity 100.0%; Pred. No. 4,2e-115;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPTIPQSAPLTPNSVMWQPLDLSRMSHSRLQHPILTLPIIDOVKTSHEVDYINDPSLAL 60
DB 1 MEPTIPQSAPLTPNSVMWQPLDLSRMSHSRLQHPILTLPIIDOVKTSHEVDYINDPSLAL 60
QY 61 TTGPKRTGAGAPELAPTPARCDQDVTHHMTSFGSRPSSVSSSTSSDRLDHPMAAPPV 120
DB 61 TTGPKRTGAGAPELAPTPARCDQDVTHHMTSFGSRPSSVSSSTSSDRLDHPMAAPPV 120
QY 121 ADOASRAVRIPKRVYHCOPLDKGPAVPELDKHFLLCEACGCKCKCKEASPTLPSCW 180
DB 121 ADOASRAVRIPKRVYHCOPLDKGPAVPELDKHFLLCEACGCKCKCKEASPTLPSCW 180
QY 181 VCNQECCLSAQTLVNTGTCMCIVGCIFFYHCTNEDEDEGSCADHPCCSRNSCCARMSFMA 240
DB 181 VCNQECCLSAQTLVNTGTCMCIVGCIFFYHCTNEDEDEGSCADHPCCSRNSCCARMSFMA 240
QY 241 LSVYLPCLLCYLPATGCVKLAQRGIDRLRRPGCKCKHTNSVICKAASGDAKTSRPDKPF 299
DB 241 LSVYLPCLLCYLPATGCVKLAQRGIDRLRRPGCKCKHTNSVICKAASGDAKTSRPDKPF 299

RESULT 3
PCT-US02-09921-1088
Sequence 1088, Application PC/TU0209921
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: DAFEO, Abel
APPLICANT: JONES, Anissa L.
APPLICANT: TRAH, Alanna-phung B.
APPLICANT: DAHL, Christopher R.
APPLICANT: GIETZEN, Darryl
APPLICANT: CHINN, Joyce
APPLICANT: DUPOUR, Gerard E.
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: YU, Jimmy Y.
APPLICANT: TUASON, Olivia
APPLICANT: YAP, Pierre E.
APPLICANT: AMSHEY, Stefan R.
APPLICANT: DAUGHERTY, Sean C.
APPLICANT: DAM, Tam C.
APPLICANT: LIU, Tommy F.
APPLICANT: NGUYEN, Duy-Viet An
APPLICANT: KLEEFELD, Yael
APPLICANT: GERSTIN JR., Edward H.

APPLICANT: PERALTA, Careyna H.
APPLICANT: DAVID, Marie H.
APPLICANT: LEWIS, Samantha A.
APPLICANT: CHEN, Alice J.
APPLICANT: PANZER, Scott R.
APPLICANT: HARRIS, Bernard
APPLICANT: FLORES, Vincent
APPLICANT: MARAUB, Rakesh
APPLICANT: LO, Audrey
APPLICANT: LAN, Ruth Y.
APPLICANT: URASHKA, Michael
TITLE OF INVENTION: SECRETORY MOLECULES
FILE REFERENCE: PT-1232 PCT
CURRENT APPLICATION NUMBER: PCT/US02/09921
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: 60/280,067; 60/280,068; 60/291,280; 60/291,849;
60/291,829; 60/299,428; 60/300,001; 60/299,776
PRIOR FILING DATE: 2001-03-29; 2001-03-29; 2001-05-16; 2001-05-17;
2001-05-17; 2001-06-19; 2001-06-20; 2001-06-20
NUMBER OF SEQ ID NOS: 1146
SOFTWARE: PERL Program
SEQ ID NO 1088
LENGTH: 375
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: LG:241742.1.orf3:2001JUN22
PCT-US02-09921-1088

Query Match 97.9%; Score 1621.5; DB 1; Length 375;
Best Local Similarity 97.7%; Pred. No. 2,2e-112;
Matches 293; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 MEPTIPQSAPLTPNSVMWQPLDLSRMSHSRLQHPILTLPIIDOVKTSHEVDYINDPSLAL 60
DB 76 MEPTIPQSAPLTPNSVMWQPLDLSRMSHSRLQHPILTLPIIDOVKTSHEVDYINDPSLAL 135
QY 61 TTGPKRTGAGAPELAPTPARCDQDVTHHMTSFGSRPSSVSSSTSSDRLDHPMAAPPV 119
DB 61 TTGPKRTGAGAPELAPTPARCDQDVTHHMTSFGSRPSSVSSSTSSDRLDHPMAAPPV 195
QY 120 VADOASRAVRIPKRVYHCOPLDKGPAVPELDKHFLLCEACGCKCKCKEASPTLPSC 179
DB 120 VADOASRAVRIPKRVYHCOPLDKGPAVPELDKHFLLCEACGCKCKCKEASPTLPSC 255
QY 180 WVCNQCCLSAQTLVNTGTCMCIVGCIFFYHCTNEDEDEGSCADHPCCSRNSCCARMSFMA 239
DB 180 WVCNQCCLSAQTLVNTGTCMCIVGCIFFYHCTNEDEDEGSCADHPCCSRNSCCARMSFMA 315
QY 240 ALSVLPCLLCYLPATGCVKLAQRGIDRLRRPGCKCKHTNSVICKAASGDAKTSRPDKPF 299
DB 316 ALSVLPCLLCYLPATGCVKLAQRGIDRLRRPGCKCKHTNSVICKAASGDAKTSRPDKPF 375

RESULT 4
US-10-082-902-5
Sequence 5, Application US/10082902
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
APPLICANT: Montgomery, Julie R.
TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
FILE REFERENCE: 28110/35878
CURRENT APPLICATION NUMBER: US/10/082,902
PRIOR FILING DATE: 2002-02-26
PRIOR APPLICATION NUMBER: US/09/370,398
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 300
TYPE: PRT

ORGANISM: Mus musculus
US-10-082-902-5

Query Match 93.8%; Score 1554.5; DB 6; Length 300;
Best Local Similarity 92.7%; Pred. No. 1.4e-107;
Matches 278; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

1 MEPP10SA-PLTPNSVVOPLDLSRMSHRIQHPILTPIDQVTSVENDYIDNPSLA 59
1 MEPPVQSSVVPNSVVOPLDLSRMSHRIQHPILTPIDQVTSVENDYIDNPSLA 60
60 LTRPKRTGAGAPLAPARCDQVTHMISFSGRPSVSSSSSTSDRLDHPAP 119
61 PATPKRPGGPELAPARCDQVTHMISFSGRPSVSSSSSTSDRLDHPAP 120
120 VADQASRAVRIOPKVHSCOPDLKGAAPPELDKHFLLCAGCKCKEASPRTPSC 179
121 VAEQASRAVRIOPKVHSCOPDLKGAAPPELDKHFLLCAGCKCKEASPRTPSC 180
180 WVCNOECISAGTIVNGTCMCTNOGIFHCNEDDESSCADHPCSSRSNCCARMSFG 239
181 WVCNOECISAGTIVNGTCMCTNOGIFHCNEDDESSCADHPCSSRSNCCARMSFG 240
240 ALSVPLCLCYLPATGCVKLAORGVDRLRRPGCRCKHTNSVICKASGDAKTSRDPKPF 299
241 ALSVPLCLCYLPATGCVKLAORGVDRLRRPGCRCKHTNSVICKASGDAKTSRDPKPF 300

RESULT 5
US-10-082-902-3
Sequence 3, Application US/10082902
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
APPLICANT: Montgomery, Julie R.
TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
FILE REFERENCE: 28110/35878
CURRENT APPLICATION NUMBER: US/10/082,902
CURRENT FILING DATE: 2002-02-26
PRIOR APPLICATION NUMBER: US/09/370,398
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 315
TYPE: PRT
ORGANISM: Homo sapiens
US-10-082-902-3

Query Match 38.7%; Score 641.5; DB 6; Length 315;
Best Local Similarity 43.6%; Pred. No. 1.9e-40;
Matches 130; Conservative 45; Mismatches 84; Indels 39; Gaps 8;

19 QPLDLSRMSHSR-----IQHPILTPIDQVTSVENDYIDNPSLAITGPK--- 65
12 QPLQTPRPGGROGPEPDRDALTQOVHVLSDQIRAIRNTNEYTGPTVPRPGIKRAP 71
66 -----RTGAGAPLAPTP-----ARCDDQVTHMISFSGRPSVSSSSSTSS 107
72 RPTQKHRLGLREHNPRLQSOVHSARAPLSRISITVSSSSRSTSTSSSS 131
108 DQRLT-DHMAPPVADQSPRAVRIOPKVHSCOPDLKGAAPPELDKHFLLCAGCKCK 166
132 EORLLGSSFSGPPVAD-----GIIVOPK-SELKPGDLK-PLSKEDGLAHAYCEDCGCK 185
167 CKECASPRTPLEPCGWCNOECISAGTIVNGTCMCTNOGIFHCNEDDESSCADHPCSC 226
186 CKECTYPRPLPSDWICDQKCLCSANMVIDTQTCVCKGFLYICSN-DDEDKADNPSCC 244
227 SRNCCARMSFGALSVPLCLCYLPATGCVKLAORGVDRLRRPGCRCKHTNSVICK 284
245 SOSHCTRMSANGVMSLPLCLMCTYLPKAGCLKLCQGCYDVRNRRPGCRCKNSNTVCK 302

RESULT 6
US-10-082-902-6
Sequence 6, Application US/10082902
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
APPLICANT: Montgomery, Julie R.
TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
FILE REFERENCE: 28110/35878
CURRENT APPLICATION NUMBER: US/10/082,902
CURRENT FILING DATE: 2002-02-26
PRIOR APPLICATION NUMBER: US/09/370,398
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 315
TYPE: PRT
ORGANISM: Mus musculus
US-10-082-902-6

Query Match 38.3%; Score 634.5; DB 6; Length 315;
Best Local Similarity 43.3%; Pred. No. 6.1e-40;
Matches 129; Conservative 46; Mismatches 84; Indels 39; Gaps 9;

19 QPLDLSRMSHSR-----IQHPILTPIDQVTSVENDYIDNPSLAITGPK--- 65
12 QPLQTPRPGGROGPEPDRDALTQOVHVLSDQIRAIRNTNEYTGPTVPRPGIKRAP 71
66 -----RTGAGAPLAPTP-----PARCDQVTHMISFSGRPSVSSSSSTSD 108
72 RPTQKHRLGLREHNPRLQSOVHSARAPLSRISITVSSSSRSTSTSSSS 131
109 QRLT-DHMAPPVADQSPRAVRIOPKVHSCOPDLKGAAPPELDKHFLLCAGCKCK 166
132 QRLGFSFSGHGAAD-----GIIVOPK-SELKPGDLK-PLSKEDGLAHAYCEDCGCK 185
167 CKECASPRTPLEPCGWCNOECISAGTIVNGTCMCTNOGIFHCNEDDESSCADHPCSC 226
186 CKECTYPRPLPSDWICDQKCLCSANMVIDTQTCVCKGFLYICSN-DDEDKADNPSCC 244
227 SRNCCARMSFGALSVPLCLCYLPATGCVKLAORGVDRLRRPGCRCKHTNSVICK 284
245 SOSHCTRMSANGVMSLPLCLMCTYLPKAGCLKLCQGCYDVRNRRPGCRCKNSNTVCK 302

RESULT 7
US-10-082-902-4
Sequence 4, Application US/10082902
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
APPLICANT: Montgomery, Julie R.
TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
FILE REFERENCE: 28110/35878
CURRENT APPLICATION NUMBER: US/10/082,902
CURRENT FILING DATE: 2002-02-26
PRIOR APPLICATION NUMBER: US/09/370,398
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 139
TYPE: PRT
ORGANISM: Homo sapiens
US-10-082-902-4

Query Match 31.9%; Score 528; DB 6; Length 139;
Best Local Similarity 62.6%; Pred. No. 1.7e-32;
Matches 87; Conservative 24; Mismatches 24; Indels 4; Gaps 1;

Query 159 CEACGCKCKEACASPTLPSCWCWNOECLCSAOTLVNVTGCMKLVGIFHYCTNEDDEGS 218
Db 1 CECGCKCKGCEGTABRTLPSCILACNRCQCLCSAESWVEYGCMLVGIYHGSNDDEDS 60
Query 219 CADHPCSCSRNCCARMSFMGALSVLPCLCYLPATGCVKLAQRGYDLRRPGRCKRH 278
Db 61 YSDNFCSCSQRNCCARMSFMGALSVLPCLCYLPATGCVKLAQRGYDLRRPGRCKRH 120
Query 279 NSYICKAAS----GDAKTS 293
Db 121 NTYCKLESQSPSRGCKPS 139
RESULT 8
US-10-082-902-7
Sequence 7 Application US/10082902
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis G.
APPLICANT: Montgomery, Julie R.
TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
FILE REFERENCE: 28110/35878
CURRENT APPLICATION NUMBER: US/10/082,902
CURRENT FILING DATE: 2002-02-26
PRIOR APPLICATION NUMBER: US/09/370,398
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 100
TYPE: PRT
ORGANISM: Homo sapiens
US-10-082-902-7

Query Match 14.3%; Score 237; DB 6; Length 100;
Best Local Similarity 39.6%; Pred. No. 2.8e-11;
Matches 44; Conservative 14; Mismatches 23; Indels 30; Gaps 3;
Query 108 DQRLDHPMAPPVADQSPRAVKIOPKVVHCQPLDLKGPVPELD----- 153
Db 5 DQRLASITPSP-----SCQSIRITQP-----GAGVHPKADGALKGEAEOAGHP 49
Query 154 -KHFILCEAGCKGCKEACASPTLPSCWCWNOECLCSAOTLVNVTGCMKLV 203
Db 50 SEHLFICEGCKGCKVPCYPTANRPLPSCHLCHQRLCSAESLIDYGTCLCV 100

RESULT 9
US-10-123-155-255
Sequence 255 Application US/10123155
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15

Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 255
LENGTH: 1432
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-155-255

Query Match 7.8%; Score 129; DB 6; Length 1432;
Best Local Similarity 26.6%; Pred. No. 0.044;
Matches 38; Conservative 10; Mismatches 71; Indels 24; Gaps 6;
Query 159 CEACGCKCKEACASPTLPSCWCWNOECLCSAOTLVNVTGCMKLVGIFHYCTNEDDEGS 218
Db 583 ctacagctcttccaatgatgcactgac--ctttgagtgcttga--acaaggaact 639
Query 219 CA-----DHPCSCSRNCCARMSFMGALSVLPCLCYLPATGCVKLAQRGYDLRR 270
Db 640 caaccagcagggctctcctaccatgagcctg-----ccagtaacagctg----- 688
Query 271 PGCRCRHTNSVICKAASGDAKTS 293
Db 689 -ac-caattgacataatggaatt 709

RESULT 10
US-10-123-155-39
Sequence 39 Application US/10123155
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 39
LENGTH: 2998
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-155-39

Query Match 7.3%; Score 121; DB 6; Length 2998;
Best Local Similarity 30.1%; Pred. No. 0.38;
Matches 40; Conservative 7; Mismatches 60; Indels 26; Gaps 7;
Query 161 ACCKCKEACASPR--TLPSWCWNOECLCSAOTLVNVTGCMKLVGIFHYCTNEDDEGS 217
Db 562 acgactctgtgaagatctctgtactatgaagt--gacacgagtc--ct----- 610
Query 218 SCADHPCSCSRNCCARMSFMGALSVLPCLCYLPATGCVKLAQRGYDLRRPGRCKRH 277
Db 611 tca---cact---ccactgagggcccgctctccctctgacataagg-----agttga 658

```
QY 278 TNSVICKAASGDA 290
      | | | |
Db 659 ttctacaanagaa 671
```

RESULT 11
US-10-123-155-185
; Sequence 185, Application US/10123155
; GENERAL INFORMATION.

ORGANISM: Homo Sapiens
US-10-123-155-185

Db 136 : | | 148
cgccaactcag

```

: APPLICANT: Stewart,Timothy A.
: APPLICANT: Tumas,Daniel
: APPLICANT: Watanabe,Colin R
: APPLICANT: Wood,William
: APPLICANT: Zhang,Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3330R1C30
: CURRENT APPLICATION NUMBER: US/10/123,155
: CURRENT FILING DATE: 2002-04-15
: Prior Application removed - See Palm or File Wrapper
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO: 113
: LENGTH: 3323
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-123-155-113

```

Query Match	6.98;	Score 114.5;	DB 6;	Length 3323;
Best Local Similarity	27.28;	Pred. No. 1.3;		
Matches 41; Conservative	7;	Mismatches 62;	Indels 41;	Gaps 9;

```

RESULT 13
US-10-123-155-331
; Sequence 331, Application US/10123155
; GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Defoige, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 331
LENGTH: 2477
TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-123-155-331

```

Query Match	6.9%;	Score 114;	DB 6;	Length 2477;
Best Local Similarity	31.3%;	Pred. No. 1;		

```

RESULT 15
PCT-US02-009944-533
: Sequence 533, Application PC/TUS0209944-533
: GENERAL INFORMATION:
: APPLICANT: INCYTE GENOMICS, INC.
: APPLICANT: DAEPO, Abel
: APPLICANT: JONES, Anisela L.
: APPLICANT: TRAN, Atlanta-phung B.
: APPLICANT: DAHL, Christopher R.
: APPLICANT: GIETZEN, Darryl

```

```

Query Match=          6.7% ; Score 111.5; DB 1; Length 633;
Best Local Similarity 21.3%; Pred No-0.35;
Matches 77; Conservative 42; Mismatches 132; Indels 111; Gaps 19.

Qy      4 PIPASALPTNSV---MVQPLDLSRMSHRLQH-----PLTLLPIDQYKTSHEVDYTDNP 56
         ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      297 PAPRPVVVTASTIRSPSYGVDPASTYSPPGANSPTPYTSSPAPA-----YRPSP 347

Qy      57 SLATITGPKRTRGGAEPLAPTRACQDDVTHIMISFSGRSPSVS----- 101
         :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      348 APATPSPVPVTVTPSPAPAYTPSPADPNPADVSVAISGPAEPASPAPRPWTIDDSFSOKFA 407

Qy      102 -SSSTSSDRLLDHMAP-----PPVADQASPRAVRI-----QPKVHCOPILDKG 145
         ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      408 PKGTSTISNQTLPRGRPATRAGPOVPRLAAGTVGMARERPASPAPRLDCHCNV--IRG 466

Qy      146 PAV-----PRELDKHFLLECAKCK-----CKEASPRLLBSCWY 181
         :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      467 PFLVAMGRSMHPPEF-----TCAYCCTSLADVCVEEDNNVNYCEKYDQFPAPLCK 518

Qy      182 CNOBCCLSAQTLVN---YGTC-MGLV-----OGIFYUOTNEDDEGSCADHPSCGSRSNC 231
         ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      519 CNTXIMGVEMHALQLQTWHHTCFVCAACKKPFGNLSLFH---MEDGEPIYCKERYIHLFEFKC 575

Qy      232 --C--ARMSENGAL-----SVLPCLLCYLPAATGCVKILAQRGYDLRRPGRCRKHNT 279
         :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      576 HGCDPFPVDAQGKFTEALGHMTDTCICAVCVHNLEG---QPEYSKDRPLCK-KHAH 629

Qy      280 SV 281
```

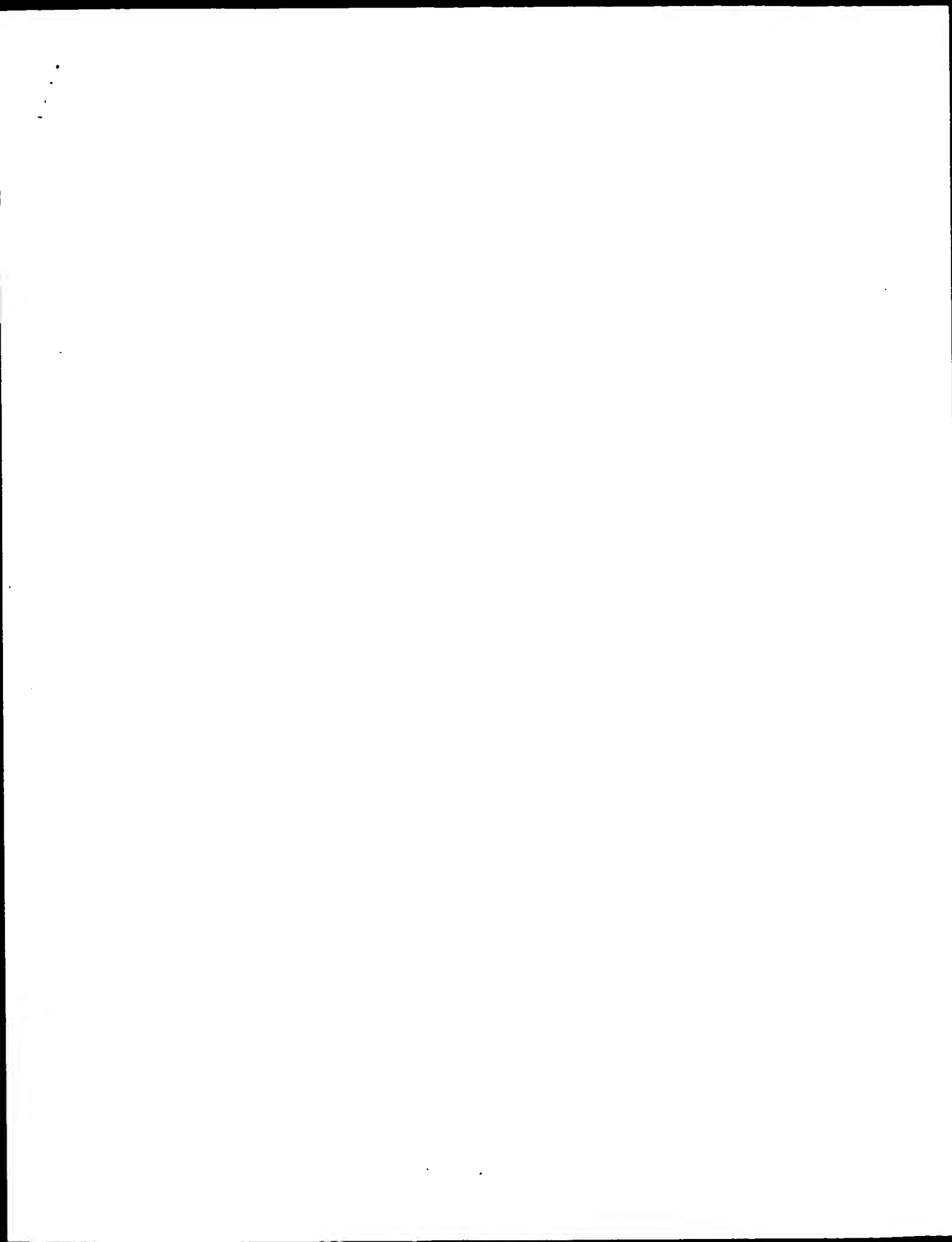

Sat May 11 12:51:25 2002

us-10-082-902-2.rapn

Page 7

Db 630 TI 631

Search completed: May 8, 2002, 10:49:05
Job time: 197 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:45:48 ; Search time 105.96 Seconds

(without alignments)
991.176 Million cell updates/sec

Title: US-10-082-902-2

Perfect score: 1657

Sequence: 1 MEPPIPQSAFLTPNSVMQV.....SVICKASGDAKTSRPDKPF 299

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3516493 seqs, 351254056 residues

Total number of hits satisfying chosen parameters: 3516493

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/2/paa/PCRTUS.COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US10.COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US083.COMB.pep.*
- 8: /cgn2_6/ptodata/2/paa/US084.COMB.pep.*
- 9: /cgn2_6/ptodata/2/paa/US085.COMB.pep.*
- 10: /cgn2_6/ptodata/2/paa/US086.COMB.pep.*
- 11: /cgn2_6/ptodata/2/paa/US087.COMB.pep.*
- 12: /cgn2_6/ptodata/2/paa/US088.COMB.pep.*
- 13: /cgn2_6/ptodata/2/paa/US089.COMB.pep.*
- 14: /cgn2_6/ptodata/2/paa/US090.COMB.pep.*
- 15: /cgn2_6/ptodata/2/paa/US091.COMB.pep.*
- 16: /cgn2_6/ptodata/2/paa/US092.COMB.pep.*
- 17: /cgn2_6/ptodata/2/paa/US093.COMB.pep.*
- 18: /cgn2_6/ptodata/2/paa/US094.COMB.pep.*
- 19: /cgn2_6/ptodata/2/paa/US095.COMB.pep.*
- 20: /cgn2_6/ptodata/2/paa/US096.COMB.pep.*
- 21: /cgn2_6/ptodata/2/paa/US097.COMB.pep.*
- 22: /cgn2_6/ptodata/2/paa/US098.COMB.pep.*
- 23: /cgn2_6/ptodata/2/paa/US099.COMB.pep.*
- 24: /cgn2_6/ptodata/2/paa/US60.COMB.pep.*
- 25: /cgn2_6/ptodata/2/paa/US60_MERGED.COMB.pep.*
- 26: /cgn2_6/ptodata/2/paa/US10_MERGED.COMB.pep.*
- 27: /cgn2_6/ptodata/2/paa/US09_MERGED.COMB.pep.*
- 28: /cgn2_6/ptodata/2/paa/US09_MERGED.COMB.pep.*
- 29: /cgn2_6/ptodata/2/paa/US08_MERGED.COMB.pep.*
- 30: /cgn2_6/ptodata/2/paa/US06_MERGED.COMB.pep.*
- 31: /cgn2_6/ptodata/2/paa/US06_MERGED.COMB.pep.*
- 32: /cgn2_6/ptodata/2/paa/PCRT_MERGED.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length DB	ID	Description
1	1657	100.0	299	US-09-370-398-2
				Sequence 2, App11

Query Match	Score 1657	DB 17	Length 299	Best Local Similarity 100.0%	Pred. No. 3.5e-125	Matches 299	Conservative 0	Mismatches 0	Gaps 0
1	MEPPIPQSAFLTPNSVMQVPLDLSRMSHSLQHPFLTLPIDQVTSRVENDYDNPISAL	60							

ALIGNMENTS

RESULT 1

US-09-370-398-2

Sequence 2, Application US/09370398

GENERAL INFORMATION:

APPLICANT: Ballinger, Dennis G.

FILE REFERENCE: 28110/35878

FILE REFERENCE: 28110/35878

CURRENT APPLICATION NUMBER: US/09/370,398

CURRENT FILING DATE: 1998-08-06

NUMBER OF SEQ ID NOS: 13

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 299

TYPE: PRT

ORGANISM: Homo sapiens

US-09-370-398-2

```

Db 1 MEPP1PQSAPLTPNSVWQPLDLSRMSHSLQHPILTLIDQVKTSHVENDYIDNPSLAL 60
Qy 61 TTGPKRTGGAPELAPTPARCDODVTHHMTSFGSRSSVSSSSSTSSDQRLDHDMAAPPV 120
Db 61 TTGPKRTGGAPELAPTPARCDODVTHHMTSFGSRSSVSSSSSTSSDQRLDHDMAAPPV 120
Qy 121 ADOASPRAVRIOPKVVHOCPLDLKGPVAPPDLKHFLCEAGCKCKEACASPTLPSCW 180
Db 121 ADOASPRAVRIOPKVVHOCPLDLKGPVAPPDLKHFLCEAGCKCKEACASPTLPSCW 180
Qy 181 VCNQECLSAOTLVNVTGCMCLVQGIFFYHCTNEDDEGSCADHPSCSRSNCCARMSFMGA 240
Db 181 VCNQECLSAOTLVNVTGCMCLVQGIFFYHCTNEDDEGSCADHPSCSRSNCCARMSFMGA 240
Qy 241 LSVVLPCLLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299
Db 241 LSVVLPCLLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299

```

RESULT 2

```

US-10-083-215-2
; Sequence 2, Application US/10083215
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
; FILE REFERENCE: 28110/35878
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: US/09/370,398
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-083-215-2

```

```

Query Match 100.0%; Score 1657; DB 26; Length 299;
Best Local Similarity 100.0%; Pred. NO. 3.5e-125;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPP1PQSAPLTPNSVWQPLDLSRMSHSLQHPILTLIDQVKTSHVENDYIDNPSLAL 60
Db 1 MEPP1PQSAPLTPNSVWQPLDLSRMSHSLQHPILTLIDQVKTSHVENDYIDNPSLAL 60
Qy 61 TTGPKRTGGAPELAPTPARCDODVTHHMTSFGSRSSVSSSSSTSSDQRLDHDMAAPPV 120
Db 61 TTGPKRTGGAPELAPTPARCDODVTHHMTSFGSRSSVSSSSSTSSDQRLDHDMAAPPV 120
Qy 121 ADOASPRAVRIOPKVVHOCPLDLKGPVAPPDLKHFLCEAGCKCKEACASPTLPSCW 180
Db 121 ADOASPRAVRIOPKVVHOCPLDLKGPVAPPDLKHFLCEAGCKCKEACASPTLPSCW 180
Qy 181 VCNQECLSAOTLVNVTGCMCLVQGIFFYHCTNEDDEGSCADHPSCSRSNCCARMSFMGA 240
Db 181 VCNQECLSAOTLVNVTGCMCLVQGIFFYHCTNEDDEGSCADHPSCSRSNCCARMSFMGA 240
Qy 241 LSVVLPCLLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299
Db 241 LSVVLPCLLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299

```

RESULT 3

```

US-10-090-190-2
; Sequence 2, Application US/10090190
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods

```

```

; FILE REFERENCE: 28110/35878
; CURRENT APPLICATION NUMBER: US/10/090,190
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US/09/370,398
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-090-190-2

```

```

Query Match 100.0%; Score 1657; DB 26; Length 299;
Best Local Similarity 100.0%; Pred. NO. 3.5e-125;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPP1PQSAPLTPNSVWQPLDLSRMSHSLQHPILTLIDQVKTSHVENDYIDNPSLAL 60
Db 1 MEPP1PQSAPLTPNSVWQPLDLSRMSHSLQHPILTLIDQVKTSHVENDYIDNPSLAL 60
Qy 61 TTGPKRTGGAPELAPTPARCDODVTHHMTSFGSRSSVSSSSSTSSDQRLDHDMAAPPV 120
Db 61 TTGPKRTGGAPELAPTPARCDODVTHHMTSFGSRSSVSSSSSTSSDQRLDHDMAAPPV 120
Qy 121 ADOASPRAVRIOPKVVHOCPLDLKGPVAPPDLKHFLCEAGCKCKEACASPTLPSCW 180
Db 121 ADOASPRAVRIOPKVVHOCPLDLKGPVAPPDLKHFLCEAGCKCKEACASPTLPSCW 180
Qy 181 VCNQECLSAOTLVNVTGCMCLVQGIFFYHCTNEDDEGSCADHPSCSRSNCCARMSFMGA 240
Db 181 VCNQECLSAOTLVNVTGCMCLVQGIFFYHCTNEDDEGSCADHPSCSRSNCCARMSFMGA 240
Qy 241 LSVVLPCLLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299
Db 241 LSVVLPCLLCYLPATGCVKLAORGYDLRRPGCRCKHTNSVICKAASGDAKTSRPDKPF 299

```

RESULT 4

```

US-09-370-398-5
; Sequence 5, Application US/09370398
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; TITLE OF INVENTION: Growth Factor Antagonist Materials and Methods
; FILE REFERENCE: 28110/35878
; CURRENT FILING DATE: US/09/370,398
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-370-398-5

```

```

Query Match 93.8%; Score 1554.5; DB 17; Length 300;
Best Local Similarity 92.7%; Pred. NO. 6.6e-117;
Matches 278; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

Qy 1 MEPP1PQSA-PLTPNSVWQPLDLSRMSHSLQHPILTLIDQVKTSHVENDYIDNPSLAL 59
Db 1 MEPP1PQSA-PLTPNSVWQPLDLSRMSHSLQHPILTLIDQVKTSHVENDYIDNPSLAL 60
Qy 60 LTTGPKRTGGAPELAPTPARCDODVTHHMTSFGSRSSVSSSSSTSSDQRLDHDMAAPPV 119
Db 60 LTTGPKRTGGAPELAPTPARCDODVTHHMTSFGSRSSVSSSSSTSSDQRLDHDMAAPPV 120
Qy 120 VAOASPRAVRIOPKVVHOCPLDLKGPVAPPDLKHFLCEAGCKCKEACASPTLPSCW 179
Db 120 VAOASPRAVRIOPKVVHOCPLDLKGPVAPPDLKHFLCEAGCKCKEACASPTLPSCW 180

```


RESULT 8
PCT-US01-27760A-819
Sequence 819, Application PC/TUS0127760A
GENERAL INFORMATION:
APPLICANT: Hysseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-115/785
CURRENT APPLICATION NUMBER: PCT/US01/27760A
CURRENT FILING DATE: 2001-10-11
PRIOR APPLICATION NUMBER: 09/687,527
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 892
SOFTWARE: Custom
SEQ ID NO 819
LENGTH: 266
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-27760A-819

Query Match 87.1%; Score 1444; DB 32; Length 266;
Best Local Similarity 97.7%; Pred. No. 4.7e-108;
Matches 239; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 35 LITLPIIDYKTSVENDYIDNPSLALTTGPKRTGAGAPELAPTPARCDODVTHHWISFSG 94
DB 2 LVTQLKLYKTSVENDYIDNPSLALTTGPKRTGAGAPELAPTPARCDODVTHHWISFSG 61
QY 95 RPSSVSSSTSSDRLDHPAPPVADQASPRAVRIQPKVHCOPLDLKGPAVPELDK 154
DB 62 RPSSVSSSTSSDRLDHPAPPVADQASPRAVRIQPKVHCOPLDLKGPAVPELDK 121
QY 155 HFLLCACGCKCKECASPRITLPSGVCNOCCLCSAOTLVNVTGCMKLVGIFHTCTND 214
DB 122 HFLLCACGCKCKECASPRITLPSGVCNOCCLCSAOTLVNVTGCMKLVGIFHTCTND 181
QY 215 DEGSACDHPSCSRNSNCARMSFGALSIVLPCLLCYLPATGCVKLAQRGYDLRRPGCR 274
DB 182 DEGSACDHPSCSRNSNCARMSFGALSIVLPCLLCYLPATGCVKLAQRGYDLRRPGCR 241
QY 275 CKHTNSYICAAASGDAKTSRPPKPF 239
DB 242 CKHTNSYICAAASGDAKTSRPPKPF 266

RESULT 9
PCT-US00-16636-9
Sequence 9, Application PC/TUS0016636
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: TUE, Henry
APPLICANT: TANG, Y. Tom
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: LAL, Pirell
APPLICANT: BANDMAN, Olga
APPLICANT: BAUGHN, Mariah R.
APPLICANT: AZIMZAI, Valda
APPLICANT: YANG, Junning
APPLICANT: REDDY, Koopa
APPLICANT: LU, Dying Anna M.
TITLE OF INVENTION: INTRACELLULAR SIGNALING MOLECULES
FILE REFERENCE: PF-0733 PCT
CURRENT APPLICATION NUMBER: PCT/US00/16636
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: 60/149,640
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/164,417
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 104
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 319

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte ID No: 2316805CD1
PCT-US00-16636-9

Query Match 41.0%; Score 679.5; DB 1; Length 319;
Best Local Similarity 43.3%; Pred. No. 3e-46; Mismatches 87; Indels 51; Gaps 9;

QY 6 PQSAPLTPNSVNV--QPLDLS--RMSHSRLQHPILTLPIIDYKTSVENDYIDNPSLALT 61
DB 3 PQNQGSSSLVVIQPSLDSRQDYEREIOTAILSDQIKAIRGSNEY-----T 54
QY 62 TGPRTKRGAPELP-----TPARCDODVTHHWISF----- 92
DB 55 EGPVYKRPAPRTAPROEKIERTHEIPIVNNVYEHRTSHLGNVLPNARGPILSR 114
QY 93 --SGRPSSVSSSTSSDRLDHPAP--PVADQASPRAVRIQPKVHCOPLDLKGPAV 149
DB 115 TSTGSAASSGSSNSASSFOGLGRSPTRVPGHRSERAIPTOPKOLIVD--DIKG-SLK 171
QY 150 PELDKHFLLCACGCKCKECASPRITLPSGVCNOCCLCSAOTLVNVTGCMKLVGIFHT 209
DB 172 EDLTKHFLLCACGCKCKECASPRITLPSCLACNRQCLCSASNVEYTCMKLVGIFHT 231
QY 210 CNEDEGSACDHPSCSRNSNCARMSFGALSIVLPCLLCYLPATGCVKLAQRGYDLR 269
DB 232 CSNDEGSYSNPSCSQSHCSRYLCMGAMSLFLPCLLCYLPAPKGLKLCRCYDWH 291
QY 270 RPGCRCKHTNSYICAAAS---GDAKTS 293
DB 292 RPGCRCKHTNSYICAAAS---GDAKTS 319

RESULT 10
US-09-488-725A-1833
Sequence 1833, Application US/09488725A
GENERAL INFORMATION:
APPLICANT: Hysseq Inc
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
FILE REFERENCE: 784FLPCT
CURRENT APPLICATION NUMBER: US/09/488,725A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US/09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US09/653,450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US09/662,191
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: US09/693,036
PRIOR FILING DATE: 2000-10-19
PRIOR APPLICATION NUMBER: US09/727,344
PRIOR FILING DATE: 2000-11-29
NUMBER OF SEQ ID NOS: 7144
SOFTWARE: pL_FL_genes_b Versions 1.0
SEQ ID NO 1833
LENGTH: 320
TYPE: PRT
ORGANISM: Homo sapiens
US-09-488-725A-1833

Query Match 41.0%; Score 679.5; DB 18; Length 319;
Best Local Similarity 43.3%; Pred. No. 3e-46;

RESULT 13
US-09-488-725A-5405
Sequence 5405, Application US/09488725A
GENERAL INFORMATION:
APPLICANT: Hyseq Inc
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
FILE REFERENCE: 784FLPCT
CURRENT APPLICATION NUMBER: US/09/488,725A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US/09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US09/653,450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US09/662,191
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: US09/693,036
PRIOR FILING DATE: 2000-10-19
PRIOR APPLICATION NUMBER: US09/727,344
PRIOR FILING DATE: 2000-11-29
NUMBER OF SEQ ID NOS: 7144
SOFTWARE: PL_genes_b Versions 1.0
SEQ ID NO 5405
LENGTH: 327
TYPE: PRT
ORGANISM: Homo sapiens
US-09-488-725A-5405

Query Match 41.0%; Score 679.5; DB 18; Length 326;

Best Local Similarity 43.3%; Pred. No. 3.1e-46; Mismatches 87; Indels 51; Gaps 9;

Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9;

QY 6 POSAPLTPNSVWV--OPLDS--RMSHSRLQHPITLPIIDOVKTSHEVDYIDNPSLALT 61
DB 10 PONQHGSSSLVVIQOPLDSRPLDYEREIOPALISLDQIKAINGSNEY-----T 61
QY 62 TGPRTKRGAPELAP-----TPARCDDVTHHWISF----- 92
DB 62 EGPVYKRPAPRTAPROEKHERTHEIIPVNNNVEHRTSHLGHAVLPNARGPILSRS 121
QY 93 --SGRPSSVSSSTSSDORLDMHAPP--PVADQSPRAVRIOPKVHCOPLDLKGPAPV 149
DB 122 TSTGSAASSGSSNASSSEGLGRPPTRPVPGHRSERAIPTOKLIYD--DLKG-SLK 178
QY 150 PEIDKHFLICEAGCKKCEKASPRILPSCWVNOECLCSAQTLYVNGTCMLVQGIYFH 209
DB 179 EDLQHKFICEQCKKCECTAPRTLPSCVACNOCCLCSAESWVEYGTOMCLVQGIYFH 238
QY 210 CTNEDDEGSCADHPSCSRSNCCARMSFGALSVLPCLCYLPATGCYKLAORGYDRLR 269
DB 239 CSNDDEGDSYDNPCSCQSOSHCSSRYLCMGAMSLFLPCLCYPAKCKLTCRCRYDWH 298
QY 270 RPGCRCKHTNSYICKAAS----GDAKTS 293
DB 299 RPGCRCKNSMTVYCKLESCPSRGGKPS 326

RESULT 14
PCT-US99-20638-7
Sequence 7, Application PC/TUS9920638
GENERAL INFORMATION:
APPLICANT: EDMONDS, Brian T.
TITLE OF INVENTION: ANTAGONISTS OF FIBROBLAST GROWTH FACTOR
FILE REFERENCE: 040902/0114
CURRENT APPLICATION NUMBER: PCT/US99/20638

CURRENT FILING DATE: 1999-09-10
EARLIER APPLICATION NUMBER: US 60/099,936
EARLIER FILING DATE: 1998-09-11
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 338
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US99-20638-7

Query Match 41.0%; Score 679.5; DB 1; Length 338;

Best Local Similarity 43.3%; Pred. No. 3.3e-46; Mismatches 87; Indels 51; Gaps 9;

Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9;

QY 6 POSAPLTPNSVWV--OPLDS--RMSHSRLQHPITLPIIDOVKTSHEVDYIDNPSLALT 61
DB 22 PONQHGSSSLVVIQOPLDSRPLDYEREIOPALISLDQIKAINGSNEY-----T 73
QY 62 TGPRTKRGAPELAP-----TPARCDDVTHHWISF----- 92
DB 74 EGPVYKRPAPRTAPROEKHERTHEIIPVNNNVEHRTSHLGHAVLPNARGPILSRS 133
QY 93 --SGRPSSVSSSTSSDORLDMHAPP--PVADQSPRAVRIOPKVHCOPLDLKGPAPV 149
DB 134 TSTGSAASSGSSNASSSEGLGRPPTRPVPGHRSERAIPTOKLIYD--DLKG-SLK 190
QY 150 PEIDKHFLICEAGCKKCEKASPRILPSCWVNOECLCSAQTLYVNGTCMLVQGIYFH 209
DB 191 EDLQHKFICEQCKKCECTAPRTLPSCVACNOCCLCSAESWVEYGTOMCLVQGIYFH 250
QY 210 CTNEDDEGSCADHPSCSRSNCCARMSFGALSVLPCLCYLPATGCYKLAORGYDRLR 269
DB 251 CSNDDEGDSYDNPCSCQSOSHCSSRYLCMGAMSLFLPCLCYPAKCKLTCRCRYDWH 310
QY 270 RPGCRCKHTNSYICKAAS----GDAKTS 293
DB 311 RPGCRCKNSMTVYCKLESCPSRGGKPS 338

RESULT 15

US-09-393-314B-7

Sequence 7, Application US/09393314B
GENERAL INFORMATION:
APPLICANT: EDMONDS, Brian T.
TITLE OF INVENTION: ANTAGONISTS OF FIBROBLAST GROWTH FACTOR
FILE REFERENCE: X-12136
CURRENT APPLICATION NUMBER: US/09/393,314B
CURRENT FILING DATE: 1998-11-09
PRIOR APPLICATION NUMBER: US 60/099,936
PRIOR FILING DATE: 1998-09-11
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 338
TYPE: PRT
ORGANISM: Homo sapiens
US-09-393-314B-7

Query Match 41.0%; Score 679.5; DB 17; Length 338;

Best Local Similarity 43.3%; Pred. No. 3.3e-46; Mismatches 87; Indels 51; Gaps 9;

Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9;

QY 6 POSAPLTPNSVWV--OPLDS--RMSHSRLQHPITLPIIDOVKTSHEVDYIDNPSLALT 61
DB 22 PONQHGSSSLVVIQOPLDSRPLDYEREIOPALISLDQIKAINGSNEY-----T 73
QY 62 TGPRTKRGAPELAP-----TPARCDDVTHHWISF----- 92
DB 74 EGPVYKRPAPRTAPROEKHERTHEIIPVNNNVEHRTSHLGHAVLPNARGPILSRS 133

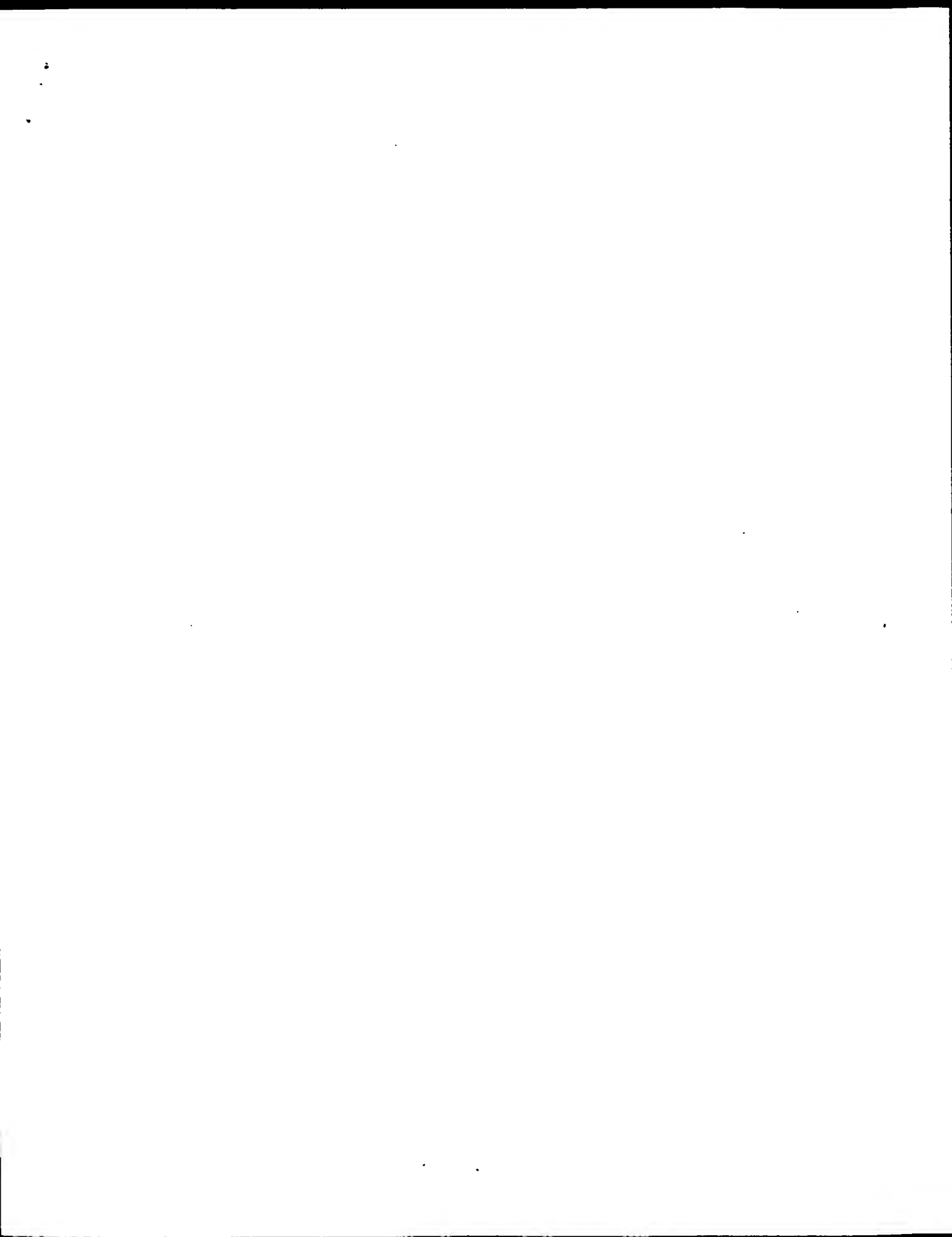
Sat May 11 12:51:24 2002

us-10-082-902-2.rapm

Page 7

[illegible]

Search completed: May 8, 2002, 10:48:30
Job time: 162 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:45:48 ; Search time 12.56 Seconds

(without alignments)
535,707 Million cell updates/sec

Title: US-10-082-902-2

Perfect score: 1657

Sequence: 1 MEPPIPQSAPLTRPSWVWP.....SVICKAASGDKTSRDKPF 299

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

212252

Total number of hits satisfying chosen parameters:

212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2.6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2.6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2.6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2.6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2.6/ptodata/2/1aa/PTCUS.COMB.pep.*
6: /cgn2.6/ptodata/2/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	650	39.2	117	3	US-08-965-903B-16
2	641.5	38.7	315	3	US-08-965-903B-8
3	528	31.9	139	3	US-08-965-903B-20
4	503.5	30.4	116	3	US-08-965-903B-9
5	493	29.8	117	3	US-08-965-903B-6
6	396	23.9	591	3	US-08-965-903B-2
7	368.5	22.2	124	3	US-08-965-903B-4
8	301	18.2	83	3	US-08-965-903B-13
9	237	14.3	100	3	US-08-965-903B-11
10	214	12.9	45	3	US-08-965-903B-18
11	202	12.2	45	3	US-08-965-903B-17
12	199	12.0	45	3	US-08-965-903B-19
13	120	7.2	2476	2	US-08-965-903B-13
14	106	6.4	721	4	US-08-872-855-7
15	105	6.4	721	4	US-08-961-392-5
16	103.5	6.2	2050	2	US-08-347-594A-2
17	102.5	6.2	577	4	US-07-728-215-29
18	99.5	6.0	578	4	US-08-981-392-13
19	98	5.9	2556	1	US-08-185-432-17
20	98	5.9	2556	1	US-08-083-590A-20
21	98	5.9	2556	1	US-08-532-384-20
22	96.5	5.8	3111	2	US-08-460-309-4
23	96.5	5.8	3111	2	US-08-125-077-4
24	95.5	5.8	2471	1	US-08-185-432-16
25	95.5	5.8	2471	1	US-08-083-590A-19
26	95.5	5.8	2471	1	US-08-532-384-19
27	95	5.7	676	3	US-08-630-172-10

28	95	5.7	676	4	US-09-375-419-10	Sequence 10, Appl
29	95	5.7	769	2	US-08-789-078-1	Sequence 1, Appl
30	95	5.7	769	2	US-08-752-633-1	Sequence 1, Appl
31	95	5.7	769	2	US-08-476-062A-45	Sequence 45, Appl
32	95	5.7	769	2	US-07-728-215-31	Sequence 31, Appl
33	95	5.7	769	5	PCT-US95-04886-1	Sequence 1, Appl
34	95	5.7	769	5	PCT-US96-01314-45	Sequence 45, Appl
35	95	5.7	1219	4	US-08-882-046-5	Sequence 2, Appl
36	95	5.7	2813	3	US-08-896-449A-2	Sequence 2, Appl
37	95	5.7	2813	3	US-09-132-652-2	Sequence 2, Appl
38	94.5	5.7	1833	3	US-08-479-722B-2	Sequence 2, Appl
39	94.5	5.7	1833	5	PCT-US95-02251-18	Sequence 18, Appl
40	94	5.7	314	2	US-08-460-309-19	Sequence 19, Appl
41	94	5.7	314	2	US-08-125-077-19	Sequence 19, Appl
42	94	5.7	2594	4	US-08-718-388-7	Sequence 9, Appl
43	94	5.7	5405	4	US-08-718-388-9	Sequence 9, Appl
44	93.5	5.6	410	6	5177197-1	Patent No. 5177197
45	93.5	5.6	1394	6	5177197-30	Patent No. 5177197

ALIGNMENTS

RESULT 1
US-08-965-903B-16
Sequence 16, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachon, Nit
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SEQUENCE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pettibory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: m-spry4 cysteine-rich region
US-08-965-903B-16

Query Match:	39.28%	Score 650;	DB 3;	Length 117;
Best Local Similarity	97.48%	Pred. No. 1,1e-51;		
Matches 114;	Conservative	0;	Mismatches 3;	Indels 0;
				Gaps 0;
Qy	183	NOEELGSAQGLVAVYGCMLGVOGIFHYCHCNEDDESSCADHSCSSRSNGCARHWSFMCALS	242	
Db	1	NOEELGSAQGLVAVYGCMLGVOGIFHYCHCNEDDESSCADHSCSSGSGNCARHSFMCALS	60	
Qy	243	VYLPCLLCITPATGCKYKLAORGYDRLRRPGCKCKTTNSVITKAAAGDCAKTSRDPKPF	299	
Db	61	VYLPCLLCITPATGCKYKLAORGYDRLRRPGCKCKTTNSVITKAAAGDCAKTSRDPKPF	117	

RESULT 2
US-08-965-903B-8
; Sequence 8, Application US/08965903B
Default No. 506037E

```

1 GENERAL INFORMATION:
2 APPLICANT: Haccohen, Nir
3 APPLICANT: Krasnow, Mark A.
4 TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
5 TITLE OF INVENTION: SEQUENCE
6 NUMBER OF SEQUENCES: 20
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Dehlinger & Associates
9 STREET: 350 Cambridge Ave., Suite 250
10 CITY: Palo Alto
11 STATE: CA
12 COUNTRY: USA
13 ZIP: 94306
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: Fastseq for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/965,903B
21 FILING DATE: 07-Nov-1997
22 CLASSIFICATION: 435
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 60/030232
25 FILING DATE: 07-Nov-1996
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Pelthoory, Joanne R
28 REGISTRATION NUMBER: 42,995
29 REFERENCE/DOCKET NUMBER: 8600-0177.30
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 650-324-0880
32 TELEFAX: 650-324-0960
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
97
```

```

Query Match      38.7% Score 641.5 DB 3 Length 315:
Best Local Similarity 43.6% Pred. No. 2.3e-50:
Matches 130: Conservative 45: Mismatches 84: Indels 39: Gaps 8

QY      19  QPLDLSRMSHR-----LQHPLEPLIPDYQKTSHEVDYIDNPSLALTTGPK---65
      ||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      12  QPLQTRDRCGRGRGEPDRDALTQQVHVLSDQLRAIRNRNINTEGTPVVPRRGLKRAP 71
      ||||: : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      66  -----KTRGAPLAPLP-----ARCQDYTNHINISFGSRPSSVSSSSSTS 107
      ||||: : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db	72	RPSTQKHREHLGLPEHQAPRLQHSQVHSSAARPLSRISTVYSSGSSSRSTSTSSSS	111
Oy	108	DOFLI...DHMAAPPVADQASPAVRITQPKVYVHCQPLDKAPVPELIDHNPFLCEAGCKCK	166
Db	132	EGALLSSSTSSSGPVAD----QIRVQPK-SELKRPGEK-PLSKEDLGLHAYRCDECKCK	185
Oy	167	CRKCAAPRLPLSPCMCNQNEGLCAQTLVYNTGCMGLVQSTFPHCNPNEDSSGADHPCSC	226
Db	186	CKEETVPRPLPSMIDCKDCKCAQSNVYIDVGTCCVKGSLFNCSN-DDEIDCAQDNPCSC	244
Oy	227	SRSNCCARMSFGALSVYPLCLCYLTPATGCVKLAQRGYDLRLPRGCKCKHTNSYICK	284
Db	245	SGSHCTKMSAGVNSLEPLPCIMCYTEPAKCKLLCGQCYDVRNRRPCCKKNSNVTYCK	302

RESULT 3
US-08-965-903B-20
; Sequence 20, Application US/08965903B

```

1 GENERAL INFORMATION:
2 APPLICANT: Hacothen, Nit
3 APPLICANT: Krasnow, Mark A.
4 TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
5 TITLE OF INVENTION: SEQUENCE
6 NUMBER OF SEQUENCES: 20
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Dehlinger & Associates
9 STREET: 350 Cambridge Ave., Suite 250
10 City: Palo Alto
11 STATE: CA
12 COUNTRY: USA
13 ZIP: 94306
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM Compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: FastSeq for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/965, 903B
21 FILING DATE: 07-NOV-1997
22 CLASSIFICATION: 435
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 60/030232
25 FILING DATE: 07-NOV-1996
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Pelthoay, Joanne R
28 REGISTRATION NUMBER: 42, 995
29 REFERENCE/DOCKET INFORMATION: 8600-0177.30
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 650-324-0880
32 TELEFAX: 650-324-0960
33 TELEX:
34 INFORMATION FOR SEQ ID NO: 20:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 139 amino acids
37 TYPE: amino acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: protein
41 FRAGMENT TYPE: internal
42 ORIGINAL SOURCE:
43 INDIVIDUAL ISOLATE: h-spry1
44 US-08-965-903B-20

```

```

Query Match:          31.9%   Score 528; DB 3; Length 139;
Best Local Similarity 62.6%   Pred. No. 1,4e-40;
Matches 87; Conservative 24; Mismatches 24; Indels 4; Gaps 1;

QY 159 CEACGKRRCKECSPTPLPSCHWCNQDELCASQTLVATGTCMKLVCQIFPHCHNEDDESS 218
      |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1  CEQGGKRRCKECSPTPLPSCLANRQGLCSKESNVEIGTQCMCLVQKIFPHCHNEDDECS 60

```

100

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pelichory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-965-903B-2

Query Match 23.9%; Score 396; DB 3; Length 591;
Best Local Similarity 29.4%; Pred. No. 7.3e-28;
Matches 95; Conservative 50; Mismatches 120; Indels 58; Gaps 10;

QY 4 PIPASAPLPNSVWQVPLDMSH-SRLQHPILPDIQVKTSHVENDYIDNPSTALT 62
DB 190 PRESERLTNEYV-----DPLQATRSQHBPAGQDNGOTTNHLILLPORNQHLQO 243
QY 63 GPRRTGAGAPLAPRARDQDYTHMISFSGRPSVSSSSSTSSD-----QRLDLM 115
DB 244 HQGHLDQO 303
QY 116 APPVADQ-ASPR-----AVRIQPKVVCQPLDK----- 144
DB 304 TKRPASKOPALPRGLKGLGLGLINQPTTKQPTATQKERNHLELLDPGAGGNG 363
QY 145 GPVAV-----PPELDKHLFLCEAGKCKEASPRTPSPCVWQNOECISAGTLVNGTCM 200
DB 364 GPLVMAGDPSLNP--IYCPRCGRRCQCSPPRLPQTWVCNKTKCLCSMESYDIASCL 421
QY 201 CLVQGIIFYHCTED---DEGS---CADHPSCSRNSCARSFMGALSVVLPCLLCYLP 253
DB 422 CCAKALFYHICARDNDLDICDNGTGPCVDNPPSCGPYKRTQRMGMGLALSTFLPCLMFYWP 481
QY 254 ATGCVKLAQGYDRLRRPCCRCK 276
DB 482 MRGCKLCEKCYGRFAGRCRCQ 504

RESULT 7
US-08-965-903B-4
Sequence 4, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachon, NIT
APPLICANT: Krasnow, Mark A.
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SEQUENCE

NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pelichory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 124 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: d-spry cysteine-rich region
US-08-965-903B-4

Query Match 22.2%; Score 368.5; DB 3; Length 124;
Best Local Similarity 46.0%; Pred. No. 3.2e-26;
Matches 57; Conservative 25; Mismatches 35; Indels 7; Gaps 2;

QY 159 CEACGCKCKEBCASPRTPSPCVWQNOECISAGTLVNGTCMCLVQGIIFYHCTED----- 214
DB 1 CPRGCRRCRCQCSPPRLPQTWVCNKTKCLCSMESYDIASCLCAKALFYHICARDNDIC 60
QY 215 DEGS---CADHPSCSRNSCARSFMGALSVVLPCLLCYLPATGCVKLAQGYDRLRRP 271
DB 61 DQNGTGPCVDNDCSGPYKRTQRMGMGLALSTFLPCLMFYWPBRGCKLCEKCYGRFAGR 120
QY 272 GCRC 275
DB 121 GCRC 124

RESULT 8
US-08-965-903B-13
Sequence 13, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachon, NIT
APPLICANT: Krasnow, Mark A.
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SEQUENCE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA

COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pelithory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 83 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: m-spryl cysteine-rich region
US-08-965-903B-13

Query Match 18.2%; Score 301; DB 3; Length 83;
Best Local Similarity 61.3%; Pred. No. 2.4e-20;
Matches 46; Conservative 16; Mismatches 13; Indels 0; Gaps 0;

QY 154 KHFLLCAGCKCKCKECSAPRTLPSCWNCNDECLCSAOTLVNVTGKCLVOGIFHYCTNE 213
DB 9 OHKFTCOCKCKCKGCGCTAPRRMPSCLCADRCCLCSASWVEYGTGKCLVGIIFYHSSND 68

QY 214 DDEGSCADHPCCSR 228
DB 69 ADGSSYSDNFCSCSQ 83

RESULT 9
US-08-965-903B-11
Sequence 11, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachosen, Nit
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SEQUENCE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pelithory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 100 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: h-spryl3 cysteine-rich region
US-08-965-903B-11

Query Match 14.3%; Score 237; DB 3; Length 100;
Best Local Similarity 39.6%; Pred. No. 1.8e-14;
Matches 44; Conservative 14; Mismatches 23; Indels 30; Gaps 3;

QY 108 DQRLDHPMAPPYADQASRAVRRIQPKVHCQPLDLGPAVPPELD----- 153
DB 5 DQRLASITPSP-----SCGSITRTQP-----GAGVHKADGALKGEAGSAGHP 49

QY 154 -KHFLCEAGCKCKCKECSAPRTLPSCWNCNDECLCSAOTLVNVTGKCLV 203
DB 50 SEHLFTCEGCKCKCYPCIAHNPSPSCWNCNDRCLCSASLIDTGLCLCV 100

RESULT 10
US-08-965-903B-18
Sequence 18, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachosen, Nit
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SEQUENCE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,903B
FILING DATE: 07-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030232
FILING DATE: 07-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pelithory, Joanne R
REGISTRATION NUMBER: 42,995
REFERENCE/DOCKET NUMBER: 8600-0177.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880

```

; TELEFAX: 650-324-0960
;
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: h-spry1 highly conserved cysteine-rich region
;
US-08-965-903B-18

Query Match          12.9%; Score 214; DB 3; Length 45;
Best Local Similarity 73.3%; Pred. No. 8e-13;
Matches 33; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 159 CEACGCKCKECASPRILPSCWNCNDECLCSAQTLYVNGTCMCLV 203
Db 1 CEACGCKCKECCTAPRILPSCWNCNDECLCSAQTLYVNGTCMCLV 45

RESULT 11
US-08-965-903B-17
; Sequence 17, Application US/08965903B
; Patent No. 6060275
;
; GENERAL INFORMATION:
; APPLICANT: Hachon, M.T.
; APPLICANT: Krasnow, Mark A.
; TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
; TITLE OF INVENTION: SEQUENCE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/965,903B
; FILING DATE: 07-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030232
; FILING DATE: 07-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Pelichory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 8600-0177.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
;
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: h-spry2 highly conserved cysteine-rich region
;
US-08-965-903B-17

Query Match          12.2%; Score 202; DB 3; Length 45;
```

```

; Best Local Similarity 66.7%; Pred. No. 9.7e-12;
; Matches 30; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 159 CEACGCKCKECASPRILPSCWNCNDECLCSAQTLYVNGTCMCLV 203
Db 1 CEACGCKCKECCTAPRILPSCWNCNDECLCSAQTLYVNGTCMCLV 45

RESULT 12
US-08-965-903B-19
; Sequence 19, Application US/08965903B
; Patent No. 6060275
;
; GENERAL INFORMATION:
; APPLICANT: Hachon, M.T.
; APPLICANT: Krasnow, Mark A.
; TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
; TITLE OF INVENTION: SEQUENCE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/965,903B
; FILING DATE: 07-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030232
; FILING DATE: 07-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Pelichory, Joanne R.
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 8600-0177.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
;
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: h-spry3 highly conserved cysteine-rich region
;
US-08-965-903B-19

Query Match          12.0%; Score 199; DB 3; Length 45;
Best Local Similarity 64.4%; Pred. No. 1.8e-11;
Matches 29; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 159 CEACGCKCKECASPRILPSCWNCNDECLCSAQTLYVNGTCMCLV 203
Db 1 CEACGCKCKECCTAPRILPSCWNCNDECLCSAQTLYVNGTCMCLV 45

RESULT 13
US-08-276-967-2
; Sequence 2, Application US/08276967
; Patent No. 5851817
;
; GENERAL INFORMATION:
; APPLICANT: Hardy, Daniel M.
; APPLICANT: Garders, David L.
```


;; TITLE OF INVENTION: Species-Specific Egg-Binding Proteins of
;; TITLE OF INVENTION: Sperm
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: Arnold, White & Durkee
;; STREET: P. O. Box 4433
;; CITY: Houston
;; STATE: Texas
;; COUNTRY: USA
;; ZIP: 77210-4433
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/276,967
;; FILING DATE: Submitted Herewith
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kitchell, Barbara S.
;; REGISTRATION NUMBER: 33,928
;; REFERENCE/DOCKET NUMBER: UTID:418KIT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 713-787-1400
;; TELEFAX: 713-789-2679
;; TELEX: 79-0924
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2476 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-276-967-2

Query Match 7.2%; Score 120; DB 2; Length 2476;
Best Local Similarity 19.8%; Pred. No. 0.036;
Matches 63; Conservative 33; Mismatches 112; Indels 110; Gaps 17;

QY 2 EPIPOSAPLPNSVYVOPLLDSRMSHSLQHL--TILPIDOVKTHVENDYIDNPSLA 59
DB 532 KPIVPTKPSPTKPTVPEEPTITPEKLTVPTEKTTPTPTKRTTPTIR--TPTIR 608
QY 60 LTGPKRTRGAPRLAPARCDQDVTHMISFSGRPSSVSSSSSTSDQDLDMAPPP 119
DB 609 TTPPERT-----TTPRTTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 654
QY 120 VADQASPAVRLOPKVY-----HCOPLDLGPVAPPELDKHFILCEAC 162
DB 655 TTPQPSPTLVPTOPAAVMPSTATVTPRTTIASCSP-----NAHF----- 696
QY 163 GKCKCK-ECASPTPLSCWY-CNOECLC-----SAQTLVNYGTCMCLVQIIFHCTNED 214
DB 697 ERCAKPVSCGSP--TPMCELFCKPVCVDPGLFSGSHCVNASSDCDYNDNYTKLGTDW 744
QY 215 DEGSCADHPESC--SRSCN-----CA----- 233
DB 755 FSPMCTEHL-CHCRSSRHECQTEKCTHTVCLKNQYGCHPYGSATCSYVDPHYLTFD 813
QY 234 --RMSFGALSYVL--PC 247
DB 814 GRRENFMKCTYILADPC 831

RESULT 14
US-08-872-855-7
; Sequence 7, Application US/08872855
; Patent No. 6121045
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean
; APPLICANT: Gearing, David

;; TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
;; TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
;; NUMBER OF SEQUENCES: 23
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
;; STREET: One Post Office Square
;; CITY: Boston
;; STATE: MA
;; COUNTRY: USA
;; ZIP: 02109-2170
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/672,855
;; FILING DATE: 11-JUN-1997
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Arnold, Beth E.
;; REGISTRATION NUMBER: 35,430
;; REFERENCE/DOCKET NUMBER: MAA-003.02
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-832-1000
;; TELEFAX: 617-832-7000
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 721 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-872-855-7

Query Match 6.4%; Score 106; DB 3; Length 721;
Best Local Similarity 24.2%; Pred. No. 0.14;
Matches 33; Conservative 13; Mismatches 79; Indels 80; Gaps 14;

QY 138 QCP---LDLGPVAPPELDKHFILCEACGCKCK-----ECASPTPLPSC----- 179
DB 214 CNPMKGLCYCTERICPLPGDEHNHGYCDKRGCKCKVGMQGRYCOECL--RYGCLHGTC 270
QY 180 ---WCV-GE---CLC-----SAQTLVNYG---TCMCLVQIIFHCTNE 213
DB 271 QQPMQCNCGQGMGGLFCMODLNYCTHHKPCENGATCTTGGGYSYCSGRGYGSNCEIE 330
QY 214 DDE-----GSCAD-----HPCSC-----SRSCARMSFGAL-----S 242
DB 331 VNECDANPKNGSGSCDLENSYCSOPGFYGNKCELSAMTCAGGCFNGRCADNPDG 390
QY 243 VVLPCLLCYLPATGCYKLAQRGYRLRRRGCRCKHT--NSYICKRASG 288
DB 391 YICCPVGY--GFNCKKTDYCSNPMCANANGACEDLNSYICQCEG 436

RESULT 15
US-08-981-392-5
; Sequence 5, Application US/08981392
; Patent No. 626025
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos Manuel Pinto
; APPLICANT: Lewis, Julian Hart
; APPLICANT: Arlaavani-Tsakonas, Spyridon
; APPLICANT: Gray, Grace
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES
; TITLE OF INVENTION: OF VERTEBRATE DELTA GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas

```

1  CITY: New York
2  STATE: NY
3  COUNTRY: USA
4  ZIP: 10036/2711
5  COMPUTER READABLE FORM:
6  MEDIUM TYPE: Diskette
7  COMPUTER: IBM Compatible
8  OPERATING SYSTEM: DOS
9  SOFTWARE: FASTSEQ version 2.0
10 CURRENT APPLICATION DATA:
11 APPLICATION NUMBER: US/08/981.3923
12 FILING DATE: 22-DEC-1997
13 CLASSIFICATION: 514
14 ATTORNEY/AGENT INFORMATION:
15 NAME: Attler, Adeline M.
16 REGISTRATION NUMBER: 32,605
17 REFERENCE/DOCKET NUMBER: 7326-038
18 TELECOMMUNICATION INFORMATION:
19 TELEPHONE: 212-790-9090
20 TELEFAX: 212-869-8864
21 TELETEX: 66111 PENNIE
22 INFORMATION FOR SEQ ID NO. 5:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 721 amino acids
25 TYPE: amino acid
26 STRANDEDNESS:
27 TOPOLOGY: unknown
28 MOLECULE TYPE: peptide
29 US-08-981-392-5

```

Query Match	6.48;	Score 106;	DB 4;	Length 721;
Best Local Similarity	23.98;	Pred. NO. 0.14;		
Matches	56;	Conservative	14;	Mismatches 70;
				Indels 94;
				Gaps 16;

```

QY 138 GCP-----LDKRPATPELDHFIPLGSAQCKCK-----ECASPRILPSC-----179
Db 214 GNPNGKGLCTEPTICLPCCDEHHGYCDKPECKCRNGWGRYCDCEI---RYPGLHGTC 270
QY 180 ---WVCN-OE---CLC-----SAQTLVNG---TCMLVGQIFVHCINE 213
Db 271 QPQWQNCQCEGNGJLFCQNDLMTYCHHKPEENATCTNNGGSYTCSCHPRGTGSNKEIE 330
QY 214 DDE-----GSCAD-----HPCSCS-----NSNCCARMSFMGALSIVYL---PCL-----248
Db 331 VNECDANPCKNKGSCSDLENSYTSCSPGFGYKNC-----ELSAMTCADGDCPFNGR 382
QY 249 -----LCVLP-----ATCGVTLAQRGIDRLRRPCRCCKHT-NSVTCKAASG 288
Db 383 CADNPDGATYCTCPVGYGFGNCEKRIIDYCSSNPCANCAREDGAGSYTQCOQEG 436

```

Search completed: May 8, 2002, 10:48:49
Job time: 181 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:45:43 ; Search time 23.43 Seconds
(without alignments)
945.280 Million cell updates/sec

Title: US-10-082-902-2
Perfect score: 1657
Sequence: 1 MERRIPQSAPLRPNVWVQD.....SVICKAASGAKTSRDKPF 299

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: A.Geneseq.1101.*
2: /SIDSB/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	679.5	41.0	319	22	AAH3688
2	679.5	41.0	319	22	AAH3688
3	679.5	41.0	319	22	AAH3688
4	672.5	40.6	319	21	AAH4074
5	668.5	40.3	319	21	AAH4074
6	654	39.5	294	21	AAH87594
7	653.5	39.4	288	21	AAH87592
8	653.5	39.4	295	21	AAH87593
9	641.5	38.7	315	19	AAH48792
10	557.5	33.6	167	21	AAH87595
11	543.5	32.8	156	21	AAH87596

12	528	31.9	139	21	AAH4184
13	493	29.8	117	19	AAH48794
14	443.5	26.8	140	19	AAH48793
15	386	23.9	591	19	AAH48796
16	291	17.6	83	21	AAH45011
17	291	17.6	83	21	AAH45012
18	237	14.3	100	19	AAH48795
19	142.5	8.6	250	21	AAH82335
20	126	7.6	738	20	AAH95322
21	121	7.3	124	21	AAH82334
22	120	7.2	2476	20	AAH67738
23	117	7.1	567	21	AAH25595
24	113.5	6.8	323	21	AAH32417
25	110.5	6.7	2813	7	AAH60053
26	110.5	6.7	2813	7	AAH60053
27	109	6.6	469	21	AAH25588
28	109	6.6	494	21	AAH25577
29	109	6.6	1686	22	AAH74944
30	109	6.6	1784	21	AAH41379
31	108	6.5	1712	13	AAH22461
32	107	6.5	464	21	AAH59440
33	107	6.5	561	21	AAH59441
34	107	6.5	582	21	AAH59442
35	107	6.5	602	21	AAH59660
36	105	6.3	363	8	AAH70437
37	105	6.3	721	21	AAH79030
38	105	6.3	1508	20	AAH27141
39	105	6.3	1508	20	AAH27141
40	105	6.3	1508	20	AAH6706
41	105	6.3	1534	19	AAH46966
42	105	6.3	1534	20	AAH27144
43	105	6.3	1534	20	AAH27144
44	105	6.3	1534	20	AAH6707
45	104.5	6.3	361	7	AAH60438

ALIGNMENTS

RESULT 1	
AAH3688	AAH3688 standard; Protein: 319 AA.
XX	XX
AC	AAH3688;
XX	XX
DE	22-Oct-2001 (first entry)
XX	XX
DE	Human polypeptide SEQ ID NO 1833.
XX	XX
KW	Human; neotrophic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW	leukemia.
XX	XX
OS	Homo sapiens.
XX	XX
FN	W020015312-A1.
XX	XX
PD	26-JUL-2001.
XX	XX
FE	26-DEC-2000; 2000MO-US34263.
XX	XX
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	19-OCT-2000; 2000US-0727344.
XX	XX

Human ORFX ORF948
Homo sapiens spron
Homo sapiens spron
Drosophila melanoga
Human secreted pro
Human secreted pro
Homo sapiens spron
Metallothionein ve
Full length pectol
Metallothionein ve
pig p105 zona pell
Protein encoded by
Human secreted pro
Sequence of von Wl
Sequence of human
Protein encoded by
TIDE protein encod
Human ADAM type me
Human ORFX ORF1143
Masking protein hl
Human deltas fragm
Human deltas fragm
Full length Human
Human Znt2 protei
Sequence of a canl
Xenopus delta prot
Human silt-1 matur
Human silt-1 matur
Protein sequence o
Amino acid sequen
Human silt-1 prote
Human silt-1 prote
Protein sequence o
Dog 10 kd alveolar

PA (HYSE-) HYSEQ INC.
XX

PI Wang J, Liu C, Asundi T, Chen R, Ma Y, Qian X, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao Q, Zhou P, Goodrich R, Dirmanac RT;
XX
WPI: 2001-442253/47.
DR
N-PSDB; AAI57844.
XX

PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
v v

the invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AA36642-AA44213) with neotropic, immunosuppressant and cyostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotides of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, activation/inhibitin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed specification.

SQ **Sequence** **319 AA;**

Query Match	41.08;	Score 679.5;	DB 22;	Length 319;
Best Local Similarity	43.38;	Score 700.0;	DB 20;	Length 319;

Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9;

QY	6	POSAPLTPSVWV--QPLTDS--ZMSISRLQHLLTLPIDOKVTSHEVNDITDNPSTALT	61
Db	3	pqnqhgsgasglvldqpalderpldyerelqptaltslsdqfkaikrgsney-----t	54
QY	62	TGPKRRRGAPLAPL-----"PRACDDVYNNHITSF-----	92
Db	55	egpsvavkrlpaptbrqekherthelirlnvnmnyehrltshlgnhavlpsnargpllsts	114
QY	93	--SGRRSVSSSSSSQOQLLDHMAP--PVADAQASPRARVQRPQVNCQPLDLSGAVP	149
Db	115	tatgaaasganssaseeqglrgsprtprpghrserritqrqpllvd--dlkg-slk	177
QY	150	PELDKFLFLCAGCKGCKCKBCASPRTPASGWNQDCLASPTLVNNTGCLVQSGFYH	209
Db	172	edtlqnlflceqgkckcgetprtlrpslclacmqdclcsaeaweygvcmlcylvqgltyn	231
QY	210	CINPEDDESGADHPCSGRSRMCARMSFMKALSVLVPLCLTLPALTQCGYKLAQNGDRLR	269
Db	232	cnndegdayadnpocsgatnserylclngamslflrplclllyppakgclclerrcywln	291

RESULT 2

ID AAB64377 standard; Protein; 319 AA.

AC AAB64377;

DT 22-MAR-2001 (first entry)

DE Amino acid sequence of human intracellular signalling molecule INTRAG,
XX

KM Human; intracellular signalling molecule; INFRA; immunosuppressive;
KM cytotoxic; neuroprotective; nocotropic; antileptodisclerotic; cancer;
KM antiinflammatory; anti-HIV; neuroleptic; antipackerial; antungual;
KM antiviral; antiparasitic; antihelminthic; antipackerial; AIDS;
KM cell proliferative disorder; arteriosclerosis autoimmune; epilepsy;
KM inflammatory disorder; Addison's disease; gastrointestinal disorder;
KM neurological disorder; Parkinson's disease; Creutzfeldt-Jakob disease;
KM mental disorder; schizophrenia; anxiety.

OS Homo sapiens

PN WO200077040-A2.
yy

PD 21-DEC-2000

16-JUN-2000; 2000WO-US166636.

PR	16-JUN-1999;	99US-01395666-
PR	17-AUG-1999;	99US-0140640-

PR 09-NOV-1999; 99US-0164417.

PA (INCY-) INCYTE GENOMICS INC.
XX

PI Yue H, Tang YT, Hillman JI

XX
DE WPT: 2001-035334 /03

DR N-PSDB; AAF32646.
 YY

PT New human intrace
PT Prevention and tr

PT inflammatory, gastrointestinal, reproductive and developmental disorders -

PS Claim 5; page 120-121; 192pp; English.

Sequence 319 AA;

Query Match	41.08;	Score 679.5;	DB 22;	Length 319;
Best Local Similarity	43.38;	Pred. No. 6e-49;		
Matches 142;	Conservative 48;	Mismatches 87;	Indels 51;	Gaps 9;

6 PQSAPLTPNSVM--OPLIDS--RMSHSRLOHPLTLPIDOVKTSHVENDYIDNPSLALT 61

Db 3 pgnqhgsgsslvlqpsldsrqrl dyereiqptal slldqika lrsney-----t 54

0Y 62 TGPKRTGGAPELAP-----TPARCDQDVTHHWISF----- 92

Db 55 egpsvkrpaprtaprqekhertheliplnvnmnyehrtshlqhavlpsnarqpllsrs 11

93 --SGRPPSSVSSSSTSSDORLLDHMAP-PVADOASPRAVRIOPKVHCOPLDLKGPVP 14

[illegible]

Db 115 tstgaasgsnssasasegllgrspptprpvrhrseratrtgpkqlivd--dlkg-slk 171
 QY 150 PELDKHRLLEACGKCKEKCASPRTPSCWNCQECISAOITVNTGTCMCIVOGIFYYH 209
 Db 172 edlqghkficeqgkckgcgcetaprllpaelacmqclcsaesmweygtcmclvkglfth 231
 QY 210 CTNEDDSCADHPCSCSRSNCCARWSEFGALSVLPCLLYLPATGCVKLAORGDRLR 269
 Db 232 csnddedgysdncpcscsqhccsrylcmgamslflpcllcypakgcklkrccrydwh 291
 QY 270 RRGCRKHTNSVICKAAS----GDAKTS 293
 Db 292 rpgcrcknsntvycklescpsrgqkps 319

RESULT 3
 AAM40474
 ID AAM40474 standard; Protein: 326 AA.
 AC AAM40474;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 5405.
 KW Human; neotrophic; immunosuppressant; cytoskeletal; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia.
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000MO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR N-PSDB: AAI59630.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 2; SEQ ID NO 5405; 10078bp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with neotrophic,
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 326 AA;

Query Match 41.0%; Score 679.5; DB 22; Length 326;
 Best Local Similarity 43.3%; Pred. No. 6,1e-49;
 Matches 142; Conservative 48; Mismatches 87; Indels 51; Gaps 9;

QY 6 POSAPLTPNSVMV--OPLDLS--RMSHSRLQHPITLPTIDQVTSHEVDINPSLAT 61
 Db 10 pnhqsgssslvlgqpsidsrpidyrelqptallslldkairgsney-----t 61
 QY 62 TGPKRTRGAPRELAP-----TPARCDOVTHMTSF----- 92
 Db 62 egpsvkvkrpaprtaqrqekherthelipnvnmehthshghavlsnarqpllsts 121
 QY 93 --SGRPSSVSSSTSSDRLDHPMAP--PVADQSPRAVRIQKVTHCOPDLKGPAPV 149
 Db 122 tstgaasgsnssasasegllgrspptprpvrhrseratrtgpkqlivd--dlkg-slk 178
 QY 150 PELDKHRLLEACGKCKEKCASPRTPSCWNCQECISAOITVNTGTCMCIVOGIFYYH 209
 Db 179 edlqghkficeqgkckgcgcetaprllpaelacmqclcsaesmweygtcmclvkglfth 238
 QY 210 CTNEDDSCADHPCSCSRSNCCARWSEFGALSVLPCLLYLPATGCVKLAORGDRLR 269
 Db 239 csnddedgysdncpcscsqhccsrylcmgamslflpcllcypakgcklkrccrydwh 298
 QY 270 RRGCRKHTNSVICKAAS----GDAKTS 293
 Db 292 rpgcrcknsntvycklescpsrgqkps 326

RESULT 4
 AAY87591
 ID AAY87591 standard; Protein: 319 AA.
 XX
 AC AAY87591;
 XX
 DT 21-JUL-2000 (first entry)
 XX
 DE Human SPROUTY-1 protein, SEQ ID NO:24.
 XX
 KW SPROUTY-1; spry1; fibroblast growth factor; FGF inhibitor;
 KW signalling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;
 KW tumour; prostate; adrenal gland.
 XX
 OS Homo sapiens.
 XX
 PI Key Location/Qualifiers
 PI FT Misc-difference 173..239
 FT /note="The base sequence encoding these residues is
 FT absent from this protein's corresponding cDNA
 FT sequence"
 XX
 PN WO200015781-A1.
 XX
 PD 23-MAR-2000.
 XX
 PR 10-SEP-1999; 99MO-US20638.
 XX
 PR 11-SEP-1998; 98US-0099936.
 XX
 PA (EUIL) LILLY & CO E.L.I.
 XX
 PI Edmonds BT;
 XX

DR WPI: 2000-271422/23.
 DR N-PSDB; AAA14071.
 XX New polynucleotide encoding a 319 amino acid sequence, used as a
 PT fibroblast growth factor (FGF) antagonist for treating prostate cancer
 PT and inhibiting angiogenesis in mammals -
 XX
 PS Claim 1: Fig 26; 93pp; English.
 XX
 CC This sequence represents a human SPROUTY-1 protein (also referred
 CC to as spry1 in the specification). SPROUTY-1 proteins (AA087591-Y87595)
 CC are antagonists of members of the fibroblast growth factor (FGF)
 CC family. The SPROUTY-1 proteins of the invention, and fragments
 CC thereof may be used in mammals to limit FGF levels in vivo and
 CC to control FGF-mediated signalling. They may be used to counteract FGF
 CC overexpression associated with certain types of tumours, and to control
 CC the growth, development or differentiation of any cell responsive to
 CC FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to
 CC control tumorigenesis, neurogenesis, metanephrogenesis, and particularly
 CC angiogenesis. They may also be used to treat conditions associated with
 CC SPROUTY-1 underexpression including polycystic kidney disease,
 CC nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
 CC insufficiency and various cancers, particularly prostatic tumours.
 CC They may additionally be used to treat a mammal with an adrenal gland
 CC pathology involving hypersensitive responses to FGF. Nucleic acid
 CC sequences encoding SPROUTY-1 proteins may be used to recombinantly
 CC express SPROUTY-1 proteins, as a source of primers and probes, and for
 CC antisense suppression of SPROUTY-1 gene expression.
 CC
 XX Sequence 319 AA:
 SQ
 Query Match 40.6%; Score 672.5; DB 21; Length 319;
 Best Local Similarity 43.0%; Pred. No. 2.3e-48;
 Matches 141; Conservative 47; Mismatches 89; Indels 51; Gaps 9;
 QY 6 POSAPLTPNSVWV--QPLDLS--RMSHSRLQHPRLTLPIDQVKTSHVENDYIDNPSLALT 61
 Db 3 pqnqhgsgsslvvlgpsldsrldyrelqpalisldqkaltgney-----t 54
 QY 62 TGPRTKRGAPRLAP-----TPARCDDVTHHWISF----- 92
 Db 55 egrsvvkrpaprtpaqekherthellplnvnnyehrlshghavlpnsargp1lars 114
 QY 93 --SGRPSSVSSSSSTSDORLLDHMAP--PVADQASPRAVRLOPKVHCQPLDLGSPVVP 149
 Db 115 tsfgsaasgsnssaseqgllygrsptlrpyrhseratltqpkqlvd--dlkg-slk 171
 QY 150 PELDKHFLLECEKCKCKEASPRTLPSGWCNDECLSAQTLVNVGTGCMCLVGGIFVH 209
 Db 172 edltqhkflcecgckckgectaprltlpsclacnrcqlcsaasmvgytcmclvkgiflyh 231
 QY 210 CTNEDDEGSCADHPSCSRSNCCARMSFGALSVLPCLLCYLPATGCKLAKORGYDLR 269
 Db 232 candegdegsadnpscsqshncsarylcmgamelflpcllcyppakgcllctrycdwln 291
 QY 270 RRGCCCKRHNSTYCKAAS---GKAITS 293
 Db 292 rpgccknntlyvcklcecpargqkps 319
 RESULT 5
 AA087594
 ID AA087594 standard; Protein; 338 AA.
 XX
 AC AA087594;
 XX
 DT 21-JUL-2000 (first entry)
 XX
 DE Human SPROUTY-1 protein, SEQ ID NO:7.
 XX
 KW SPROUTY-1; spry1; fibroblast growth factor; FGF inhibitor;
 signalling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;

KW tumour; prostate; adrenal gland.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 191..258
 FT /note="The base sequence encoding these residues is
 FT absent from this protein's corresponding cDNA
 FT sequence"
 XX
 XX WO200015781-A1.
 XX
 XX 23-MAR-2000.
 XX
 XX 10-SEP-1999; 99WO-US20638.
 XX
 XX 11-SEP-1998; 98US-0099936.
 XX
 XX (ELIL) LILLY & CO ELI.
 XX
 XX Edmonds ET;
 XX
 XX WPI: 2000-271422/23.
 XX N-PSDB; AAA14074.
 XX
 PT New polynucleotide encoding a 319 amino acid sequence, used as a
 PT fibroblast growth factor (FGF) antagonist for treating prostate cancer
 PT and inhibiting angiogenesis in mammals -
 XX
 PS Example 4; Fig 7; 93pp; English.
 XX
 CC This sequence represents a human SPROUTY-1 protein (also referred
 CC to as spry1 in the specification). SPROUTY-1 proteins (AA087591-Y87595)
 CC are antagonists of members of the fibroblast growth factor (FGF)
 CC family. The SPROUTY-1 proteins of the invention, and fragments
 CC thereof may be used in mammals to limit FGF levels in vivo and
 CC to control FGF-mediated signalling. They may be used to counteract FGF
 CC overexpression associated with certain types of tumours, and to control
 CC the growth, development or differentiation of any cell responsive to
 CC FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to
 CC control tumorigenesis, neurogenesis, metanephrogenesis, and particularly
 CC angiogenesis. They may also be used to treat conditions associated with
 CC SPROUTY-1 underexpression including polycystic kidney disease,
 CC nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
 CC insufficiency and various cancers, particularly prostatic tumours.
 CC They may additionally be used to treat a mammal with an adrenal gland
 CC pathology involving hypersensitive responses to FGF. Nucleic acid
 CC sequences encoding SPROUTY-1 proteins may be used to recombinantly
 CC express SPROUTY-1 proteins, as a source of primers and probes, and for
 CC antisense suppression of SPROUTY-1 gene expression.
 CC
 XX Sequence 338 AA:
 SQ
 Query Match 40.3%; Score 668.5; DB 21; Length 338;
 Best Local Similarity 43.0%; Pred. No. 5.3e-48;
 Matches 141; Conservative 48; Mismatches 88; Indels 51; Gaps 9;
 QY 6 POSAPLTPNSVWV--QPLDLS--RMSHSRLQHPRLTLPIDQVKTSHVENDYIDNPSLALT 61
 Db 22 pqnqhgsgsslvvlgpsldsrldyrelqpalisldqkaltgney-----t 73
 QY 62 TGPRTKRGAPRLAP-----TPARCDDVTHHWISF----- 92
 Db 74 egrsvvkrpaprtpaqekherthellplnvnnyehrlshghavlpnsargp1lars 133
 QY 93 --SGRPSSVSSSSSTSDORLLDHMAP--PVADQASPRAVRLOPKVHCQPLDLGSPVVP 149
 Db 134 tsfgsaasgsnssaseqgllygrsptlrpyrhseratltqpkqlvd--dlkg-slk 190
 QY 150 PELDKHFLLECEKCKCKEASPRTLPSGWCNDECLSAQTLVNVGTGCMCLVGGIFVH 209
 Db 191 edltqhkflcecgckckgectaprltlpsclacnrcqlcsaasmvgytcmclvkgiflyh 250

QY 210 CTNEDDEGSCADHFCGSCSRNCCARMSFMCALSVLPCLLCYLPATGCGVKLAORGDRRLR 269
 Db 251 csnddedgdsysdncpscsgshcserylcmgamslflpcclycpakgciklccrcydwih 310
 QY 270 RFGCRCKHTNSVICKAAS---GDAAKT 293
 Db 311 ifgcrccknsntvcklescpsrgqkps 338

RESULT 6

AAV87592
 ID AAV87592 standard; Protein; 294 AA.

AAV87592;
 21-JUL-2000 (first entry)

Human SPROUTY-1 protein, SEQ ID NO:3.

SPROUTY-1; spry1; fibroblast growth factor; FGF inhibitor;
 signalling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;
 tumour; prostate; adrenal gland.

Homo sapiens.

WO200015781-A1.

23-MAR-2000.

10-SEP-1999; 99WO-US20638.

11-SEP-1998; 98US-0099936.

(ELIL) LILLY & CO ELI.

Edmonds BT;

WPI: 2000-271422/23.

N-PSDB; AAA14072.

New polynucleotide encoding a 319 amino acid sequence, used as a
 fibroblast growth factor (FGF) antagonist for treating prostate cancer
 and inhibiting angiogenesis in mammals -

Disclosure: Fig 3; 93pp; English.

This sequence represents a human SPROUTY-1 protein (also referred
 to as spry1 in the specification). SPROUTY-1 proteins (AAV87591-Y87595)
 are antagonists of members of the fibroblast growth factor (FGF)
 family. The SPROUTY-1 proteins of the invention, and fragments
 thereof may be used in mammals to limit FGF levels in vivo and
 to control FGF-mediated signalling. They may be used to counteract FGF
 overexpression associated with certain types of tumours, and to control
 the growth, development or differentiation of any cell responsive to
 FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to
 control tumorigenesis, neurogenesis, metanephrogenesis, and particularly
 angiogenesis. They may also be used to treat conditions associated with
 SPROUTY-1 underexpression including polycystic kidney disease,
 nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
 insufficiency and various cancers, particularly prostatic tumours.
 They may additionally be used to treat a mammal with an adrenal gland
 pathology involving hypersensitive responses to FGF. Nucleic acid
 sequences encoding SPROUTY-1 proteins may be used to recombinantly
 express SPROUTY-1 proteins, as a source of primers and probes, and for
 antisense suppression of SPROUTY-1 gene expression.

Sequence 294 AA;

Query Match 39.58; Score 654; DB 21; Length 294;
 Best Local Similarity 44.28; Pred. No. 7.3e-47;
 Matches 133; Conservative 42; Mismatches 78; Indels 48; Gaps 8;

QY 30 RLDG-HPLTILPIDOVKTSHEVDYIDNPFLATGTGPKRRCGAPELAP-----T 77
 Db 5 rldgpralislqldkairgsney-----tegpsvvrkraprtaprqekherthell 56
 QY 78 PARCDQDVTHHWISF-----SGRPSSVSSSTSSDQRLDHMAP 117
 Db 57 plannnyehrhshghavlpnargpjlrsrstsagsanssaseqglgrsp 116
 QY 118 P-PVADQSPRAVRIDPKVHVCOPLDKGPAPVPELDHFLFCEACGCKCKEASPTL 176
 Db 117 trpyprghseralrtgpkqlvq--dlkg-slkedltqhkflfcccqgckkccgetprtl 173
 QY 177 PSCWVCNQDECLCSAQTLVNTGTCMKLVGCIFFHCTNEDDEGSCADHPCSCSRNCCARMS 236
 Db 174 psclacnrcqclcsesemweyglcmclvkglfthcsndedgdsdncpscsgshcseryl 233
 QY 237 FMGALSVLPCLLCYLPATGCGVKLAORGDRRLRRPFCRCCKHTNSVICKAAS---GDAKT 292
 Db 234 cmgamslflpcclycpakgciklccrcydwihprgcrccknsntvcklescpsrgqkps 293
 QY 293 S 293
 Db 294 s 294

RESULT 7

AAV87597
 ID AAV87597 standard; Protein; 288 AA.

AAV87597;
 21-JUL-2000 (first entry)

Human SPROUTY protein long consensus sequence.

SPROUTY; spry; fibroblast growth factor; FGF inhibitor;

signalling; tumorigenesis; neurogenesis; metanephrogenesis; cancer;
 tumour; prostate; adrenal gland; consensus sequence.

Homo sapiens.

WO200015781-A1.

23-MAR-2000.

10-SEP-1999; 99WO-US20638.

11-SEP-1998; 98US-0099936.

(ELIL) LILLY & CO ELI.

Edmonds BT;

WPI: 2000-271422/23.

New polynucleotide encoding a 319 amino acid sequence, used as a
 fibroblast growth factor (FGF) antagonist for treating prostate cancer
 and inhibiting angiogenesis in mammals -

Disclosure: Fig 2; 93pp; English.

Sequences AAV87596-Y87597 represent, respectively, short and long
 SPROUTY protein consensus sequences. The invention relates to human
 SPROUTY-1 proteins (also referred to as spry1 in the specification;
 CC AAV87591- AAV87595) which are antagonists of members of the fibroblast
 growth factor (FGF) family. The SPROUTY-1 proteins of the invention, and
 fragments thereof may be used in mammals to limit FGF levels in vivo and
 to control FGF-mediated signalling. They may be used to counteract FGF
 overexpression associated with certain types of tumours, and to control
 the growth, development or differentiation of any cell responsive to
 FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to
 control tumorigenesis, neurogenesis, metanephrogenesis, and particularly

KM tumour; prostate; adrenal gland; consensus sequence.
 XX Homo sapiens.
 OS
 XX WO200015781-A1.
 PN
 XX 23-MAR-2000.
 PD
 XX 10-SEP-1999; 99MO-US20638.
 PF
 XX 11-SEP-1998; 98US-0099936.
 PR
 XX (ELIL) LILLY & CO ELI.
 PA
 XX Edmonds BT;
 PI
 XX Edmonds BT;
 DR WPI: 2000-271422/23.
 XX
 PT New polynucleotide encoding a 319 amino acid sequence, used as a
 PT fibroblast growth factor (FGF) antagonist for treating prostate cancer
 PT and inhibiting angiogenesis in mammals -
 XX
 PS Disclosure: Fig 1; 93pp; English.
 XX
 CC Sequences AAY87596-Y87597 represent, respectively, short and long
 CC SPROUTY protein consensus sequences. The invention relates to human
 CC SPROUTY-1 proteins (also referred to as spry) in the specification;
 CC AAY87591-AAY87595) which are antagonists of members of the fibroblast
 CC growth factor (FGF) family. The SPROUTY-1 proteins of the invention, and
 CC fragments thereof may be used in mammals to limit FGF levels in vivo and
 CC to control FGF-mediated signalling. They may be used to counteract FGF
 CC overexpression associated with certain types of tumours, and to control
 CC the growth, development or differentiation of any cell responsive to
 CC FGF-mediated signalling. For example, SPROUTY-1 proteins may be used to
 CC control tumorigenesis, neurogenesis, metanephrogenesis, and particularly
 CC angiogenesis. They may also be used to treat conditions associated with
 CC SPROUTY-1 underexpression including polycystic kidney disease,
 CC nephropathy, lung fibrosis, hepatic fibrogenesis, endometrial
 CC insufficiency and various cancers, particularly prostatic tumours.
 CC They may additionally be used to treat a mammal with an adrenal gland
 CC pathology involving hypersensitive responses to FGF. Nucleic acid
 CC sequences encoding SPROUTY-1 proteins may be used to recombinantly
 CC express SPROUTY-1 proteins, as a source of primers and probes, and for
 CC antisense suppression of SPROUTY-1 gene expression.
 CC
 XX Sequence 156 AA;
 SO
 Query Match 32.8%; Score 543.5; DB 21; Length 156;
 Best Local Similarity 59.0%; Pred. No. 5.9e-38;
 Matches 92; Conservative 29; Mismatches 30; Indels 5; Gaps 2;
 OY 142 DKGAVAPPELDKHLCEGCKKCKEASPTLTSCHWVNCNOELCSQTLVNTGTCMC 201
 DB 2 dkgv-sleddtqkhlcecgkckecaptlpsclancrqcscsewveygtmc 60
 OY 202 LVQGFHYCTNEDEGSCADHPCSCSRNSCCARMSFGALSVLPCLLCYLPATGCVLA 261
 DB 61 lvggfhyctnedegscadhpccsrnsccarmsfgalsvlpcllcytpatgcvla 120
 OY 262 ORGYRLRPGCRCKKHTNSVYCKKAS----GDAKTS 233
 DB 121 rrcydwlrpgcrckkhtnsvyckkas----gda kts 156
 RESULT 12
 AAB41184
 ID AAB41184 standard; Protein: 139 AA.
 AC AAB41184;
 XX
 DT 08-FEB-2001 (first entry)
 XX

DE Human ORFX ORF948 polypeptide sequence SEQ ID NO:1896.
 XX
 XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
 KW vulnery; antiparacitic; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasoregic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antineoplastic;
 KW antiviral; antibacterial; antifungal; antineumatic; antihypertensive;
 KW antineumatic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antineoplastic disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 XX Homo sapiens.
 OS
 XX WO200058473-A2.
 PN
 XX 05-OCT-2000.
 PD
 XX 31-MAR-2000; 2000MO-US08621.
 PF
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shinkets RA, Leach M;
 PI WPI: 2000-602362/57.
 DR N-PSDB; AAC75393.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 11; Page 1453; 5507pp; English.
 PS
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosolic; hepatotropic; vulnery;
 CC antiparacitic; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasoregic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antineoplastic; antibacterial; antiviral; antifungal; antineumatic;
 CC antihypertensive; and antineumatic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antineoplastic disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX Sequence 139 AA;
 SO
 Query Match 31.9%; Score 528; DB 21; Length 139;
 Best Local Similarity 62.6%; Pred. No. 1e-36;
 Matches 87; Conservative 24; Mismatches 24; Indels 4; Gaps 1;
 OY 159 CEAQCKCKEASPTLTSCHWVNCNOELCSQTLVNTGTCMLVQGFHYCTNEDEGS 218

QY 219 CADHPCSCSRNCCARMSFNGALSVLPCLLCTLPATGCVKLA-----ORGYD----- 266
 DB 61 ysdnpscscsqhscsrylcmgamslfllpcllyppakgclksgyvmtgslagadvtlp 120
 QY 267 -----RLRRPGCRCKH 277
 DB 121 tlslywtrazppgyrvrvnh 138

RESULT 15

AAW48796
 ID AAW48796 standard; Protein; 591 AA.

AC AAW48796;

DT 11-SEP-1998 (first entry)

DE Drosophila melanogaster sprouty protein.

KW spry; sprouty protein.

OS Drosophila melanogaster.

PN MO9820032-A1.

PD 14-MAY-1998.

PF 07-NOV-1997; 97WO-US20398.

PR 07-NOV-1996; 96US-0030232.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Hachsen N, Krasnow MA;

DR WPI; 1998-286662/25.

DR N-PSDB; AAV32402.

PT New isolated human sprouty polypeptide(s) - used to develop products for treating disorders involving angiogenic dysfunction, e.g.

PT cancers, diabetic retinopathy, glaucoma, psoriasis or inflammation

PS Example 5; Fig 5C; 90pp; English.

CC The sequence is that of Drosophila sprouty protein which was used

CC in the isolation of human spry genes.

CC Sequence 591 AA;

Query Match 23.9%; Score 396; DB 19; Length 591;
 Best Local Similarity 29.4%; Pred. No. 5.8e-25;
 Matches 95; Conservative 50; Mismatches 120; Indels 58; Gaps 10;

QY 4 PTFQSAPLTPNSVWQPLRDSRMSH-SRLQHPRLTLPIDQVKTSHVENDYINDPSLATLT 62
 DB 190 ppreseetlney-----dtrlqhatrsqhpqgqngqtcthhlllpqtrnhllhlg 243
 QY 63 GPKRTTGGARPELAPTRACDQDVTNNHISFGRRSSVSSSSSTSSD-----QRLDMM 115
 DB 244 hqqlhqqgqqgqqgqqgqqlhqqgqqghatlatctqatcavgsdhtgllhshlqnst 303
 QY 116 APPPVADO-ASPR-----AVRIQPKVYHCOPLDLK----- 144
 DB 304 tkppakqprlprlmglglglglnprlltkqpratqkermhaeelqpggagng 363
 QY 145 GRPV-----PRELKHFLCEAGCKKCEKASPRTLPSCWVNOECTCSAQTLYVYTCM 200
 DB 364 gplvmagdpallnp--lvprcgrcgcgcqsprrlpqtlwcnkctclcsaesvldyascl 421
 QY 201 CLVQGFYFHTCTED---DEGS---CADHPCSCSRNCCARMSFNGALSVLPCLLCTLP 253
 DB 422 cckalrfyhcarnldcdcdgngcpcvndpscscgpykrtqrgwlgalsiflpcwlfywp 481

QY 254 ATGVKLAORGYDRRLRRPGCRCK 276
 DB 482 mrgcmklcekygrfagrgctcg 504

Search completed: May 8, 2002, 10:46:15
 Job time: 32 sec